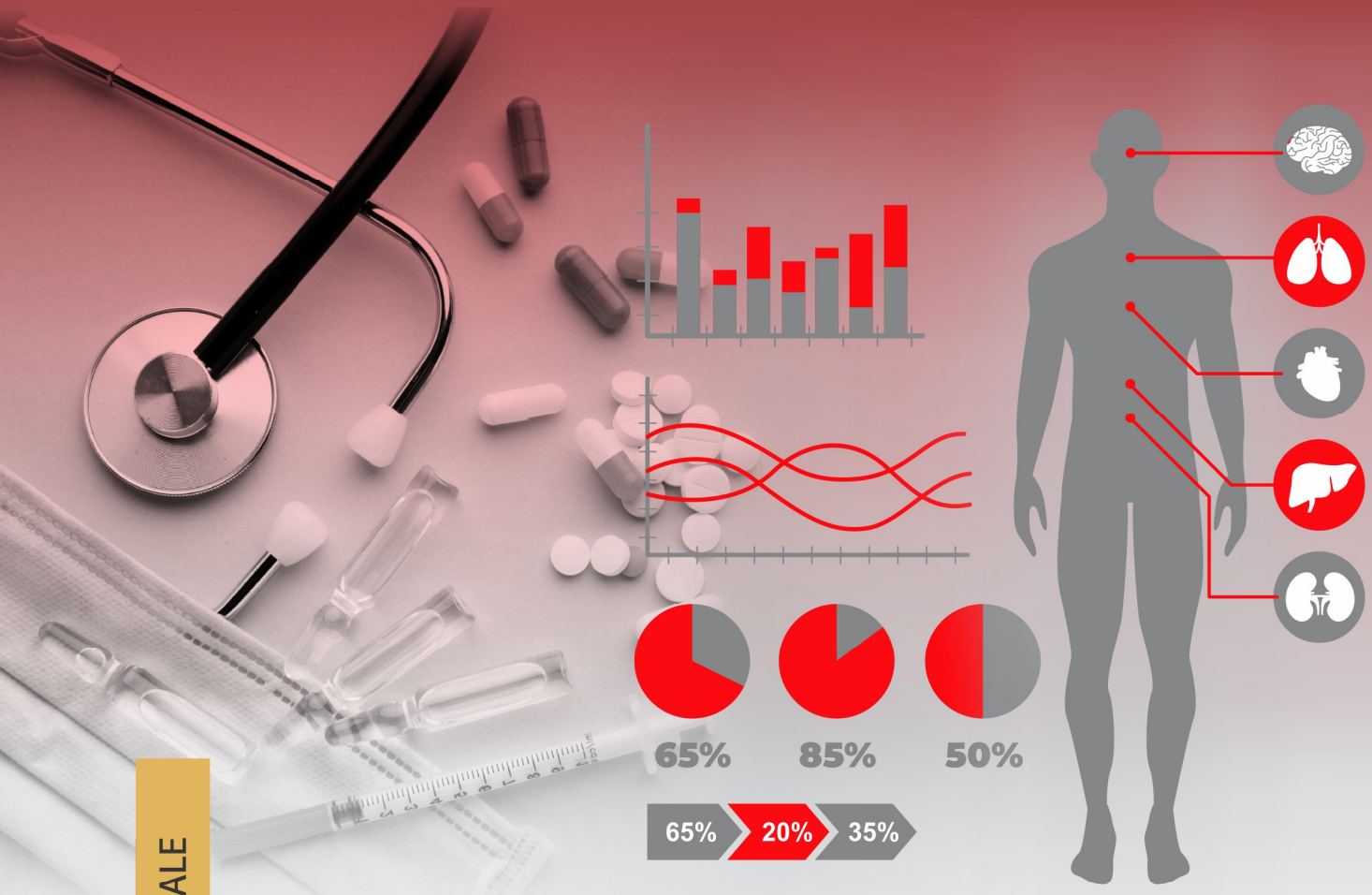
UNIVERSITÉ
DE MÉDICINE ET PHARMACIE
„VICTOR BABEŞ" DE TIMIŞOARA

CĂLIN MUNTEAN

INFORMATIQUE MÉDICALE

Méthodes, systèmes et pratique. Horizon 2026

Cours et travaux pratiques pour les étudiants en médecine



MANUALE

Éditions „Victor Babeş"
Timişoara, 2026

Editura „Victor Babeş”
Piața Eftimie Murgu nr. 2, cam. 316, 300041 Timişoara
Tel./Fax 0256 495 210
e-mail: evb@umft.ro
https://www.umft.ro/ro/organizare-evb/

Director: Prof. univ. dr. Sorin Ursoniu

Colecția: MANUALE
Coordonatori colecție: **Prof. univ. dr. Codruţa Şoica**
Prof. univ. dr. Daniel Lighezan

Referent științific: Prof. univ. dr. Daliborca Vlad

Indicativ CNCSIS: 324

© 2026
Toate drepturile asupra acestei ediții sunt rezervate. Reproducerea parțială sau integrală a textului, pe orice suport, fără acordul scris al autorilor este interzisă și se va sancționa conform legilor în vigoare.

ISBN 978-606-786-643-8

Les exemples de patients présentés dans cet ouvrage sont fictifs. Les informations relatives à la législation, aux plateformes nationales et aux standards reflètent la situation au 27 septembre 2026. L'ouvrage ne remplace pas les recommandations cliniques et les réglementations en vigueur, qui doivent être consultées dans leur version actualisée.

Table des matières

Préface ix
Comment lire cet ouvrage x
Liste des abréviations xii
Partie I. Cours 1
Chapitre 1. L'informatique médicale : objet, domaines, compétences 2
1.1 Définitions et objet 2
1.2 Repères historiques : des premiers ordinateurs à l'hôpital à l'intelligence artificielle générative 4
1.3 Les domaines de l'informatique biomédicale : des molécules aux populations 6
1.4 Les compétences numériques du médecin 8
1.5 La documentation scientifique 9
1.6 L'informatique médicale en Roumanie et à Timişoara 12
1.7 Comment utiliser cet ouvrage 13
Chapitre 2. Données, information et connaissance en médecine 16
2.1 Les données médicales 16
2.2 La théorie de l'information appliquée à la médecine 18
2.3 Données, information, connaissance, sagesse 22
2.4 La communication de l'information dans l'acte médical 24
2.5 La médecine fondée sur les preuves au format numérique 26
2.6 L'information en biologie 28
Chapitre 3. Données structurées, fichiers et bases de données 31
3.1 Fichiers et formats 31
3.2 La modélisation des données 32
3.3 Systèmes de gestion de bases de données relationnelles et SQL 34
3.4 Bases NoSQL, entrepôts et lacs de données, mégadonnées 35
3.5 Modèles communs de données pour la recherche 37
3.6 La qualité des données et les registres médicaux 40
3.7 Les outils de collecte électronique des données 41
Chapitre 4. Classifications, nomenclatures et terminologies médicales 45
4.1 Pourquoi coder 45
4.2 La CIM-10 et les modifications nationales 46
4.3 La CIM-11 49
4.4 Autres classifications de l'OMS 52

4.5 Terminologies cliniques 52
4.6 Les groupes homogènes de malades (DRG) 54
4.7 Le codage assisté par ordinateur et l'intelligence artificielle 56
Chapitre 5. Éthique, protection des données et sécurité de l'information en santé 59
5.1 Principes éthiques appliqués aux données de santé 59
5.2 Le RGPD pour le médecin 61
5.3 L'espace européen des données de santé 64
5.4 La sécurité de l'information 65
5.5 Équité et biais algorithmique 68
5.6 Intégrité académique 69
Chapitre 6. Signaux biomédicaux : acquisition, traitement et interprétation assistée 73
6.1 Le signal biomédical 73
6.2 La chaîne d'acquisition 75
6.3 Le traitement numérique 77
6.4 Artefacts et qualité du signal 81
6.5 Formats et bases de données ouvertes 82
6.6 L'interprétation assistée par l'intelligence artificielle 83
6.7 La situation en Roumanie en 2026 85
Chapitre 7. Imagerie médicale numérique : DICOM, PACS et analyse assistée 88
7.1 L'image numérique 88
7.2 Les modalités d'imagerie du point de vue informatique 90
7.3 La compression des images 91
7.4 Le standard DICOM 92
7.5 PACS, RIS, VNA et flux de travail radiologique 94
7.6 Traitement et analyse des images 96
7.7 L'intelligence artificielle en imagerie, 2026 97
Chapitre 8. Le dossier de santé électronique et l'interopérabilité 102
8.1 Le dossier électronique : concepts et effets 102
8.2 L'interopérabilité 105
8.3 Le résumé international du patient et les services transfrontaliers 110
8.4 L'espace européen des données de santé 111
8.5 La Roumanie en 2026 112
8.6 Utilisabilité, sécurité et facteurs humains 114
Chapitre 9. La télémédecine : formes, preuves, limites et cadre juridique 117
9.1 Définitions et taxonomie 117

9.2 L'infrastructure technique 119
9.3 Preuves d'efficacité 2020-2026 121
9.4 Limites et risques 124
9.5 Le cadre juridique et le remboursement 126
9.6 La documentation de la téléconsultation 127
Chapitre 10. La santé mobile, les dispositifs portables et les modèles de langage en consultation 131
10.1 La santé mobile 131
10.2 Dispositifs portables et capteurs 133
10.3 Le logiciel dispositif médical 136
10.4 L'interopérabilité des dispositifs 139
10.5 Les grands modèles de langage en consultation 140
10.6 La situation en Roumanie en 2026 143
Chapitre 11. Le système d'information, la sécurité de l'information et la signature électronique 146
11.1 Le système d'information et le système informatique 146
11.2 Le management de la sécurité de l'information 149
11.3 Notions de cryptographie pour le médecin 151
11.4 La signature électronique 153
11.5 La normalisation en informatique médicale 155
11.6 Le facteur humain 157
Chapitre 12. Les systèmes d'information hospitaliers et le système informatique unique intégré de l'assurance maladie 161
12.1 Le système d'information de l'hôpital : composantes et modules 161
12.2 La prescription informatisée à l'hôpital et l'aide à la décision clinique 164
12.3 Étude de cas : l'informatisation de la réanimation 165
12.4 Le système informatique de l'assurance maladie en Roumanie 167
12.5 La transformation numérique des hôpitaux et la maturité numérique 170
12.6 Indicateurs et tableaux de bord 171
12.7 La résilience : quand le système s'arrête 172
Chapitre 13. L'intelligence artificielle en médecine : méthodes, évaluation, réglementation (2026) 175
13.1 Qu'est-ce que l'intelligence artificielle en médecine ? 175
13.2 Le cycle de vie d'un modèle 178
13.3 L'évaluation de la performance 179
13.4 Les preuves cliniques 2020-2026 183
13.5 L'IA explicable 185

13.6 Les risques 185
13.7 La réglementation en 2026 186
13.8 Note de renvoi 188
Chapitre 14. Synthèse : les compétences du médecin dans l'environnement numérique et l'éthique de l'utilisation de l'intelligence artificielle 191
14.1 La carte de la discipline 191
14.2 Les compétences à la fin de la première année 192
14.3 Le calendrier des réglementations 2024-2031 qui changeront la pratique 194
14.4 L'éthique de l'utilisation de l'IA par l'étudiant et le médecin 196
14.5 Orientations après 2026 198
14.6 Questions de révision 199
Partie II. Travaux pratiques 206
Travaux pratiques 1. L'ordinateur, le système d'exploitation et l'espace de travail collaboratif (Google Workspace) 207
L1.1 L'ordinateur et les fichiers : notions minimales 207
L1.2 Partage, versions et sécurité du compte 209
Travaux pratiques 2. L'information médicale sur le web : recherche et évaluation de la qualité. 213
L2.1 De la question clinique à la requête 213
L2.2 Wikipédia : un point de départ, pas une source à citer 215
L2.3 Outils d'évaluation de l'information destinée aux patients 215
L2.4 La réponse d'un chatbot comparée au guide 216
Travaux pratiques 3. Formulaires électroniques de collecte des données : Google Forms et Epi Info 219
L3.1 Le protocole et le dictionnaire de données 220
L3.2 Construction du formulaire dans Google Forms 221
L3.3 Le test pilote 222
L3.4 La collecte et l'export 222
L3.5 Le nettoyage des données dans Google Sheets 222
L3.6 Le tableau descriptif 223
L3.7 La variante Epi Info 7 223
L3.8 La protection des données 224
Travaux pratiques 4. Codage des diagnostics avec les outils en ligne de l'OMS (CIM-10 et CIM-11) 226
L4.1 Préparation (5 minutes) 227
L4.2 La liste tabulaire de la CIM-10 (15 minutes) 227
L4.3 Le navigateur CIM-11 et le Coding Tool (15 minutes) 228

L4.4 Codage de dix cas (35 minutes) 228
L4.5 Exercice de postcoordination (10 minutes) 229
L4.6 La correspondance CIM-10-CIM-11 (10 minutes) 229
L4.7 La cause initiale du décès et l'ordre dans le cluster (10 minutes) 229
Travaux pratiques 5. Citer correctement, prévenir le plagiat et rédiger un essai scientifique 232
L5.1 Anatomie d'une référence Vancouver 233
L5.2 Le gestionnaire de références 234
L5.3 Paraphrase, citation, copie 235
L5.4 Le rapport de similitude 236
L5.5 Déclarer l'utilisation de l'IA générative 236
L5.6 La structure de l'essai 236
Travaux pratiques 6. Analyse des signaux ECG et EEG avec des logiciels libres 238
Travaux pratiques 7. Traitement des images médicales avec ImageJ/Fiji 244
L7.1 Installation et préparation 245
L7.2 Ouverture de l'image et lecture de l'en-tête DICOM 246
L7.3 Histogramme et fenêtre de visualisation 246
L7.4 Calibration spatiale et mesures 246
L7.5 Filtrage du bruit 247
L7.6 Segmentation par seuillage et comptage des objets 247
L7.7 Enregistrement d'une macro et exportation des résultats 248
L7.8 Discussion : limites et reproductibilité 248
Travaux pratiques 8. Travailler avec un dossier de santé électronique : démo OpenEMR et ressources FHIR 250
L8.1 Partie A : le patient fictif dans OpenEMR 251
L8.2 Partie B : les mêmes concepts sous forme de ressources FHIR 252
L8.3 Partie C : comparaison avec la prescription électronique en Roumanie 254
Travaux pratiques 9. Simulation d'une téléconsultation : préparation, consentement, documentation et limites 256
Travaux pratiques 10. La consultation simulée avec un modèle de langage : vérification, erreurs et déclaration 262
Travaux pratiques 11. Sécurité pratique : mots de passe, authentification multifacteur, vérification d'une signature électronique et reconnaissance de l'hameçonnage 268
L11.1 La robustesse d'un mot de passe (20 minutes) 269
L11.2 La validation en deux étapes sur le compte Google (15 minutes) 269
L11.3 Signature et vérification d'un PDF fictif dans Acrobat Reader (35 minutes) 270
L11.4 Le certificat TLS d'un site (10 minutes) 271

L11.5 Reconnaître l'hameçonnage : test avec des courriels fictifs (25 minutes)..... 271
L11.6 La procédure de signalement d'un incident (10 minutes) 272
Travaux pratiques 12. Le flux d'information au cabinet, en polyclinique et à l'hôpital : modélisation et indicateurs 274
L12.1 Le scénario 275
L12.2 Le diagramme en couloirs dans draw.io 276
L12.3 Documents et éléments de données 276
L12.4 Le jeu de données et le tableur 277
L12.5 Le calcul des indicateurs 278
L12.6 Interprétation 279
Travaux pratiques 13. L'évaluation critique d'un outil d'IA : de la matrice de confusion au débat 281
Travaux pratiques 14. L'examen pratique : modèle de sujet, barème et révision..... 287
L14.1 Le sujet type 288
L14.2 Corrigés et barème 290
L14.3 Erreurs fréquentes 291
L14.4 Plan de préparation 291
Annexe A. Réponses aux questions d'autoévaluation 293
Annexe B. Glossaire 294
Références bibliographiques 300

Préface

Pourquoi un ouvrage d'informatique médicale daté aussi précisément, "Horizon 2026" ? Parce que, contrairement à l'anatomie, l'objet de cette discipline se déplace sous nos yeux. Pendant les années où j'ai écrit ces pages, une partie des repères enseignés aux étudiants ont changé deux fois. La onzième révision de la classification des maladies de l'OMS, la CIM-11, est en vigueur depuis 2022, et la fiche d'information de l'OMS de mai 2026 recense 132 pays à différents stades de mise en œuvre. À la date où j'achève cet ouvrage, les hôpitaux roumains regroupent pourtant les cas à partir de diagnostics codés en ICD-10-AM, troisième édition. Le standard HL7 FHIR est enseigné dans sa version R4, sur laquelle repose la majorité des implémentations, bien que la R5 soit publiée depuis 2023 et que la R6 soit soumise au vote. Le règlement européen sur l'intelligence artificielle a été adopté en 2024, et en juillet 2026 le calendrier des obligations applicables aux systèmes à haut risque a été décalé par le règlement (UE) 2026/1744. L'espace européen des données de santé est en vigueur depuis mars 2025, mais ses obligations s'appliquent de manière échelonnée jusqu'en 2031. En Roumanie, le portail eSănătateaMea a été lancé le 1er septembre 2026, et la plateforme informatique de l'assurance maladie a été modernisée dans le cadre du PNRR, le plan national de relance et de résilience.

Un étudiant de première année exercera dans un système où toutes ces échéances seront déjà dépassées. Il ne m'a pas semblé honnête de lui offrir une image sans date. J'ai préféré présenter chaque fait roumain et européen comme un fait en mouvement, avec la date à laquelle il a été vérifié et sa source officielle, afin que le lecteur puisse refaire la vérification dans un an ou dans cinq.

L'intelligence artificielle occupe dans ce livre plus de place que dans les manuels d'il y a dix ans, mais pas la place centrale. J'ai essayé de la présenter telle qu'un jeune médecin la rencontre : comme un instrument qu'il faut évaluer avec les mêmes questions qu'un test diagnostique, qui peut se tromper de façon convaincante et pour lequel la responsabilité reste celle de l'humain. C'est pourquoi le calcul de la sensibilité et de la valeur prédictive, la vérification des références et la protection des données du patient apparaissent dans plusieurs chapitres, et pas seulement dans celui consacré à l'intelligence artificielle. Un étudiant qui maîtrise ces notions saura aussi juger les outils qui apparaîtront après 2026.

L'ouvrage ne remplace pas les manuels de la discipline, mais les cite et les présuppose. Le manuel d'informatique médicale et de biostatistique de Mihalaș et Lungeanu [1] reste la référence pour la biostatistique et pour les fondements que ce livre ne reprend pas en détail. L'ouvrage de Focșa sur l'informatique médicale, des données à l'intelligence artificielle [2], et le guide de Mihalaș sur l'intelligence artificielle en médecine [3] couvrent, sous d'autres angles, bon nombre des thèmes abordés ici. Je recommande de les lire en parallèle. Ce qu'apporte le présent ouvrage, c'est le lien étroit avec le programme de 2025, avec les travaux pratiques de chaque semaine et avec la situation de la Roumanie à une date précise.

La structure suit le programme analytique de 2025 de la Discipline d'informatique médicale et de biostatistique de l'Université de médecine et de pharmacie "Victor Babeș" de Timișoara : 14 cours et 14 séances de travaux pratiques, soit une paire par semaine. Le même contenu est enseigné aux étudiants des filières en roumain, en anglais et en français, c'est pourquoi les termes techniques apparaissent aussi en anglais lors de leur première occurrence. Le premier chapitre définit la discipline et les compétences numériques du médecin. Viennent ensuite les données et la

théorie de l'information, les bases de données, les classifications et les terminologies, l'éthique et la protection des données. Les chapitres 6 et 7 traitent des signaux et des images, les chapitres 8-10 du dossier électronique, de la télémédecine et de la santé mobile, et les chapitres 11-13 du système d'information, des systèmes hospitaliers et nationaux et de l'intelligence artificielle. Le dernier chapitre est une synthèse, avec la carte de la discipline, le calendrier des réglementations et les règles d'utilisation de l'intelligence artificielle.

Les travaux pratiques ont une histoire plus longue que ce livre. Certains protocoles proviennent des volumes d'applications pratiques de la discipline dont j'ai été coauteur : le guide de travaux pratiques de 2010 [4], le recueil d'applications pratiques d'informatique médicale et de biostatistique paru la même année [5] et le volume de 2018 consacré au traitement des données biomédicales [6]. Ils sont repris et actualisés pour les outils gratuits de 2026 : Google Forms et Google Sheets, les navigateurs de l'OMS pour la CIM, Fiji, OpenEMR, des serveurs FHIR de test, des bases publiques de signaux. D'autres sont nouveaux, parce que leurs thèmes n'existaient pas dans le programme il y a dix ans. La consultation simulée avec un modèle de langage et l'évaluation critique d'un outil d'intelligence artificielle en sont deux exemples.

Chaque chapitre de cours est construit de la même façon. Il commence par des objectifs opérationnels, formulés comme des actions que l'étudiant peut accomplir à la fin. Le contenu va des notions aux méthodes et, le cas échéant, à la situation en Roumanie, présentée dans un encadré séparé. Suivent un encadré consacré aux pièges fréquents, c'est-à-dire aux erreurs que j'ai vues se répéter en séminaire et à l'examen, puis un encadré qui rassemble les idées à retenir. Le chapitre se termine par cinq questions à choix multiples, au format du concours de l'internat (rezidențiat roumain), dont les réponses se trouvent à l'annexe A. Chaque séance de travaux pratiques comporte un objectif, les outils avec leur version et leur date d'accès, les étapes de travail, le travail à rendre avec un barème indicatif et une variante pour les salles sans réseau.

Je me suis tenu à quelques règles de travail, que j'exige aussi des étudiants. Chaque référence a été effectivement trouvée dans PubMed, dans CrossRef ou sur la page de l'institution émettrice avant d'être citée, et les chiffres sont repris tels qu'ils figurent dans la source. Les graphiques fondés sur des données réelles nomment leur source. Ceux qui ont été construits à des fins didactiques sont signalés comme données synthétiques. Les patients des exemples sont fictifs.

Les informations sur la législation roumaine et européenne, sur les plateformes nationales et sur les versions des standards reflètent la situation au 27 septembre 2026. Certaines changeront avant que l'ouvrage ne parvienne au lecteur. Lorsqu'une source officielle n'a pas pu être trouvée, j'ai préféré dire que je ne l'avais pas trouvée plutôt que de combler le vide par une supposition.

Comment lire cet ouvrage

L'ouvrage peut être lu dans l'ordre, semaine après semaine, mais ce n'est pas obligatoire. Les chapitres renvoient les uns aux autres, et la carte de la discipline présentée au chapitre 14 montre comment ils s'articulent. Qui a besoin rapidement d'un thème (par exemple, le codage d'un diagnostic ou le calcul de la valeur prédictive positive) peut partir directement du chapitre concerné et de la séance de travaux pratiques correspondante.

Pour la préparation courante, je recommande trois étapes. On lit d'abord les objectifs, pour savoir ce que l'on doit être capable de faire à la fin. On parcourt ensuite le texte avec les figures et les tableaux, qui ne sont pas des illustrations mais font partie du contenu. Enfin, on répond aux questions sans regarder les réponses de l'annexe A, et l'on ne vérifie qu'ensuite. Les pièges

fréquents méritent d'être lus deux fois, une fois avant la séance de travaux pratiques et une fois avant l'examen.

Les travaux pratiques se font devant un ordinateur allumé. Ils demandent un compte Google et quelques logiciels gratuits, mais chacun comporte aussi une variante hors ligne. Les fichiers rendus deviennent le matériel de l'examen pratique, dont le modèle complet se trouve dans le TP 14.

Les citations sont numérotées selon le style Vancouver, dans l'ordre de leur première apparition, et la liste complète se trouve à la fin de l'ouvrage. Les abréviations sont réunies dans une liste séparée, et les principaux termes sont brièvement définis dans le glossaire de l'annexe B, avec un renvoi au chapitre où ils sont expliqués.

Une dernière recommandation concerne les données. Toute échéance légale, toute version de standard et tout chiffre relatif aux plateformes roumaines figurant dans ce livre doivent être vérifiés à la source avant d'être utilisés dans un travail ou dans une décision. L'ouvrage indique chaque fois la source. Je n'ai rien écrit sur des patients réels, et le lecteur est prié d'appliquer la même règle dans ses exercices, y compris lorsqu'il utilise un outil d'intelligence artificielle.

Timișoara, septembre 2026 — Călin Muntean

Liste des abréviations

Liste des abréviations utilisées dans l'ouvrage, par ordre alphabétique. Pour les termes empruntés à l'anglais, la dénomination originale figure entre parenthèses.

Abréviation	Forme développée
ACHI	liste australienne des interventions de santé (angl. *Australian Classification of Health Interventions*)
ACID	atomicité, cohérence, isolation, durabilité (propriétés des transactions)
ACOM	modèle de contenu pour les observations de santé défini dans IEEE 11073-10206
ACR	Collège américain de radiologie (angl. *American College of Radiology*)
ADC	convertisseur analogique-numérique (angl. *analog-to-digital converter*) ; voir CAN
ADN	acide désoxyribonucléique
ADR	Autoritatea pentru Digitalizarea României (Autorité pour la numérisation de la Roumanie)
ADT	module de gestion administrative des patients : admission, sortie, transfert (angl. *admission, discharge, transfer*)
AES	norme de chiffrement symétrique (angl. *Advanced Encryption Standard*)
AI Act	règlement (UE) 2024/1689 sur l'intelligence artificielle
AIPD	analyse d'impact relative à la protection des données (angl. *data protection impact assessment*, DPIA)
AMIA	Association américaine d'informatique médicale (angl. *American Medical Informatics Association*)
ANMDMR	Agenția Națională a Medicamentului și a Dispozitivelor Medicale din România (Agence nationale du médicament et des dispositifs médicaux de Roumanie)
ANSPDCP	Autoritatea Națională de Supraveghere a Prelucrării Datelor cu Caracter Personal (Autorité nationale de surveillance du traitement des données à caractère personnel, Roumanie)
API	interface de programmation d'applications (angl. *application programming interface*)
ASCII	code américain normalisé pour l'échange d'information (angl. *American Standard Code for Information Interchange*)
ASRO	Asociația de Standardizare din România (Association de normalisation de Roumanie)
ATC	classification anatomique, thérapeutique et chimique des médicaments
ATI	anestezie și terapie intensivă (anesthésie-réanimation, Roumanie)
AUC	aire sous la courbe ROC (angl. *area under the curve*)
AVC	accident vasculaire cérébral
BI-RADS	système de compte rendu et de données en imagerie mammaire (angl. *Breast Imaging Reporting and Data System*)
BMI	informatique biomédicale (angl. *biomedical informatics*)
BPCO	bronchopneumopathie chronique obstructive
BPMN	notation de modélisation des processus (angl. *Business Process Model and Notation*)
CA	autorité de certification (angl. *certification authority*)
CAN	convertisseur analogique-numérique
CAS	casa de asigurări de sănătate (caisse départementale d'assurance maladie, Roumanie)
CDA	architecture des documents cliniques (angl. *Clinical Document Architecture*)
CDC	Centres pour le contrôle et la prévention des maladies, États-Unis (angl. *Centers for Disease Control and Prevention*)
CDS	aide à la décision clinique (angl. *clinical decision support*)
CE	marquage de conformité européenne
CEAS	cardul electronic de asigurări de sănătate (carte électronique d'assurance maladie, Roumanie)

Abréviation	Forme développée
CEI	cartea electronică de identitate (carte d'identité électronique, Roumanie)
CEN	Comité européen de normalisation (CEN/TC 251 : comité technique pour l'informatique médicale)
CGM	mesure continue du glucose (angl. *continuous glucose monitoring*)
CIF	Classification internationale du fonctionnement, du handicap et de la santé (angl. *International Classification of Functioning, Disability and Health*, ICF)
CIM	Classification internationale des maladies (angl. *International Classification of Diseases*, ICD) ; CIM-10, CIM-11
CLAIM	lignes directrices pour la publication des études d'IA en imagerie médicale
CMD	catégorie majeure de diagnostic
CNAS	Casa Națională de Asigurări de Sănătate (Caisse nationale d'assurance maladie, Roumanie)
CNP	cod numeric personal (code numérique personnel, Roumanie)
CONSORT-AI	lignes directrices pour la publication des essais randomisés portant sur des interventions d'IA
COPE	Comité d'éthique des publications (angl. *Committee on Publication Ethics*)
CPOE	prescription informatisée (angl. *computerized provider order entry*)
CR	radiographie à plaques phosphore lues ultérieurement (angl. *computed radiography*)
CRF	cahier d'observation (angl. *case report form*)
CSIRT	équipe de réponse aux incidents de sécurité informatique (angl. *Computer Security Incident Response Team*)
CSV	valeurs séparées par des virgules (angl. *comma-separated values*)
DBT	tomosynthèse mammaire numérique (angl. *digital breast tomosynthesis*)
DDD	dose définie journalière (angl. *defined daily dose*)
DECIDE-AI	lignes directrices pour la publication de l'évaluation clinique précoce des systèmes d'aide à la décision fondés sur l'IA
DECODE	référentiel des compétences en santé numérique dans l'enseignement médical (angl. *Digital Health Competencies in Medical Education*)
DES	Dosarul electronic de sănătate (dossier électronique de santé du patient, Roumanie)
DICOM	standard pour l'imagerie et les communications numériques en médecine (angl. *Digital Imaging and Communications in Medicine*)
DiGA	applications de santé numériques remboursées en Allemagne (all. *Digitale Gesundheitsanwendungen*)
DigComp	cadre européen des compétences numériques pour les citoyens
DIKW	données, information, connaissance, sagesse (angl. *data, information, knowledge, wisdom*)
DMS	durée moyenne de séjour
DNSC	Directoratul Național de Securitate Cibernetică (Direction nationale de la cybersécurité, Roumanie)
DOI	identifiant numérique d'objet (angl. *Digital Object Identifier*)
DPO	délégué à la protection des données (angl. *data protection officer*)
DR	radiographie numérique directe (angl. *digital radiography*)
DRG	groupes homogènes de malades (angl. *diagnosis-related groups*)
DSE, EHR	dossier de santé électronique (angl. *electronic health record*)
E2EE	chiffrement de bout en bout (angl. *end-to-end encryption*)
EBM	médecine fondée sur les preuves (angl. *evidence-based medicine*)
ECG	électrocardiogramme
ECR, RCT	essai contrôlé randomisé (angl. *randomized controlled trial*)
EDC	saisie électronique des données (angl. *electronic data capture*)
EDF, EDF+	format européen de données pour les signaux (angl. *European Data Format*)

Abréviation	Forme développée
EEG	électroencéphalogramme
EEHRxF	format européen d'échange des dossiers de santé électroniques (angl. *European Electronic Health Record Exchange Format*)
EFMI	Fédération européenne d'informatique médicale (angl. *European Federation for Medical Informatics*)
EHDEN	réseau européen de données de santé et de preuves (angl. *European Health Data & Evidence Network*)
EHDS	espace européen des données de santé (angl. *European Health Data Space*), règlement (UE) 2025/327
eIDAS	règlement (UE) n° 910/2014 sur l'identification électronique et les services de confiance (eIDAS 2 : règlement (UE) 2024/1183)
EMG	électromyogramme
EMR	dossier médical électronique (angl. *electronic medical record*)
EMRAM	modèle d'adoption du dossier médical électronique (angl. *Electronic Medical Record Adoption Model*)
ENISA	Agence de l'Union européenne pour la cybersécurité
ERAS	réhabilitation améliorée après chirurgie (angl. *Enhanced Recovery After Surgery*)
ERP	progiciel de gestion intégré (angl. *enterprise resource planning*)
ETL	extraction, transformation, chargement (angl. *extract, transform, load*)
EUDI	portefeuille européen d'identité numérique (angl. *European Digital Identity Wallet*)
FA	fibrillation atriale
FAIR	faciles à trouver, accessibles, interopérables, réutilisables (angl. *Findable, Accessible, Interoperable, Reusable*)
FC	fréquence cardiaque
FDA	Agence américaine des produits alimentaires et des médicaments (angl. *Food and Drug Administration*)
FFT	transformée de Fourier rapide (angl. *fast Fourier transform*)
FHIR	ressources rapides pour l'interopérabilité en santé (angl. *Fast Healthcare Interoperability Resources*), standard HL7
FIPS	normes fédérales américaines de traitement de l'information (angl. *Federal Information Processing Standards*)
FN	faux négatif
FOCG	foaia de observație clinică generală (dossier d'observation clinique générale, Roumanie)
FP	faux positif
GHS, GHSP	profil Bluetooth générique pour les capteurs de santé (angl. *Generic Health Sensor Profile*)
GRADE	système d'évaluation de la certitude des preuves et de la force des recommandations (angl. *Grading of Recommendations Assessment, Development and Evaluation*)
HbA1c	hémoglobine glyquée
HG	Hotărârea Guvernului (décision du Gouvernement, Roumanie)
HL7	organisation et famille de standards Health Level Seven (HL7 v2 : le standard de messagerie)
HTA	hypertension artérielle
HTTPS	protocole HTTP sécurisé par TLS
IA	intelligence artificielle (angl. *artificial intelligence*, AI)
IBSI	initiative de standardisation des biomarqueurs d'imagerie (angl. *Image Biomarker Standardisation Initiative*)
IC	intervalle de confiance
ICD	sigle anglais de la Classification internationale des maladies ; voir CIM
ICD-10-AM	modification australienne de la CIM-10, utilisée en Roumanie dans sa troisième édition

Abréviation	Forme développée
ICD-10-CM	modification clinique de la CIM-10, utilisée aux États-Unis
ICHI	Classification internationale des interventions de santé
ICM	indice de case-mix
ICMJE	Comité international des rédacteurs de revues médicales (angl. *International Committee of Medical Journal Editors*)
IDMP	normes ISO d'identification des médicaments (angl. *Identification of Medicinal Products*)
IEC	Commission électrotechnique internationale (angl. *International Electrotechnical Commission*)
IEEE	Institut des ingénieurs électriciens et électroniciens (angl. *Institute of Electrical and Electronics Engineers*)
IHE	initiative d'intégration des systèmes des établissements de santé (angl. *Integrating the Healthcare Enterprise*)
IMIA	Association internationale d'informatique médicale (angl. *International Medical Informatics Association*)
INMSS	Institutul Național de Management al Serviciilor de Sănătate (Institut national de management des services de santé, Roumanie)
IOD	définition d'un objet d'information DICOM (angl. *Information Object Definition*)
I-PASS	protocole de transmission du patient : gravité, résumé, actions, situations imprévues, synthèse par le receveur
IPS	résumé international du patient (angl. *International Patient Summary*)
IRM	imagerie par résonance magnétique (angl. *magnetic resonance imaging*, MRI)
ISMS	système de management de la sécurité de l'information (angl. *information security management system*)
ISO	Organisation internationale de normalisation (ISO/TC 215 : comité technique pour l'informatique médicale)
ITIL	bibliothèque de bonnes pratiques pour la gestion des services informatiques (angl. *IT Infrastructure Library*)
IVDR	règlement (UE) 2017/746 relatif aux dispositifs médicaux de diagnostic in vitro
JSON	notation des objets JavaScript (angl. *JavaScript Object Notation*)
LIS	système d'information de laboratoire (angl. *laboratory information system*)
LLM	grand modèle de langage (angl. *large language model*)
LOINC	noms et codes des identifiants d'observations (angl. *Logical Observation Identifiers Names and Codes*)
MARS	échelle d'évaluation des applications mobiles (angl. *Mobile App Rating Scale*)
MDCG	Groupe de coordination en matière de dispositifs médicaux (angl. *Medical Device Coordination Group*)
MDR	règlement (UE) 2017/745 relatif aux dispositifs médicaux
MeSH	vocabulaire contrôlé de la NLM pour les sujets médicaux (angl. *Medical Subject Headings*)
MFA	authentification multifacteur (angl. *multi-factor authentication*)
ML	apprentissage automatique (angl. *machine learning*)
MMS	linéarisation de la CIM-11 pour les statistiques de mortalité et de morbidité
MPPS	notification de l'étape de procédure réalisée par la modalité (angl. *Modality Performed Procedure Step*)
MS	Ministerul Sănătății (ministère de la Santé, Roumanie)
NAM	Académie nationale de médecine des États-Unis (angl. *National Academy of Medicine*)
NIS2	directive (UE) 2022/2555 concernant la sécurité des réseaux et des systèmes d'information
NIST	Institut national des normes et de la technologie, États-Unis (angl. *National Institute of Standards and Technology*)
NLM	Bibliothèque nationale de médecine des États-Unis (angl. *National Library of Medicine*)

Abréviation	Forme développée
NNT	nombre de sujets à traiter (angl. *number needed to treat*)
NoSQL	bases de données non relationnelles (angl. *not only SQL*)
OHDSI	réseau scientifique pour les données de santé observationnelles (angl. *Observational Health Data Sciences and Informatics*)
OMOP CDM	modèle de données commun OMOP (angl. *Observational Medical Outcomes Partnership Common Data Model*)
OMS	Organisation mondiale de la santé
OR	rapport de cotes (angl. *odds ratio*)
OUG	ordonanță de urgență a Guvernului (ordonnance d'urgence du Gouvernement, Roumanie)
PA	pression artérielle
PACS	système d'archivage et de transmission d'images (angl. *Picture Archiving and Communication System*)
PDMS	système de gestion des données du patient en soins intensifs (angl. *patient data management system*)
PGHD	données de santé générées par le patient (angl. *patient-generated health data*)
PHR	dossier personnel de santé (angl. *personal health record*)
PIAS	Platforma informatică din asigurările de sănătate (plateforme informatique de l'assurance maladie, Roumanie)
PICO	population, intervention, comparateur, critère de jugement (angl. *Population, Intervention, Comparator, Outcome*)
PKI	infrastructure à clés publiques (angl. *public key infrastructure*)
PMID	identifiant d'un article dans PubMed
PNRR	Planul Național de Redresare și Reziliență (plan national de relance et de résilience, Roumanie)
PPG	photopléthysmographie (angl. *photoplethysmography*)
QCM	question à choix multiple, avec plusieurs réponses correctes (type de question)
QCS	question à choix simple, avec une seule réponse correcte (type de question)
RAG	génération augmentée par récupération (angl. *retrieval-augmented generation*)
RAM	mémoire vive de l'ordinateur (angl. *random access memory*)
RBAC	contrôle d'accès fondé sur les rôles (angl. *role-based access control*)
REST	style d'interface web utilisé par FHIR et DICOMweb (angl. *representational state transfer*)
RGPD, GDPR	règlement général sur la protection des données, règlement (UE) 2016/679 (angl. *General Data Protection Regulation*)
RIM	modèle d'information de référence de HL7 (angl. *Reference Information Model*)
RIS	système d'information de radiologie (angl. *Radiology Information System*)
RMN	résonance magnétique nucléaire
RMSSD	racine carrée de la moyenne des carrés des différences successives entre les intervalles NN
ROC	caractéristique de fonctionnement du récepteur, courbe ROC (angl. *receiver operating characteristic*)
ROI	région d'intérêt (angl. *region of interest*)
RR	intervalle entre deux ondes R successives ; dans le contexte des études, risque relatif
RSA	algorithme de chiffrement asymétrique Rivest-Shamir-Adleman
SBAR	protocole de communication : situation, contexte, évaluation, recommandation (angl. *Situation, Background, Assessment, Recommendation*)
SDNN	écart-type des intervalles NN
SGBD	système de gestion de bases de données (angl. *database management system*)
SHAP	explications additives fondées sur les valeurs de Shapley (angl. *SHapley Additive exPlanations*)

Abréviation	Forme développée
SIEC-ATI	sistemul informatic pentru evidența clinică a secțiilor ATI (système informatique d'enregistrement clinique des services d'anesthésie-réanimation, Roumanie)
SIH, HIS	système d'information hospitalier (angl. *hospital information system*)
SIPE	Sistemul informatic de prescripție electronică (système informatique de prescription électronique, Roumanie)
SIUI	Sistemul informatic unic integrat al asigurărilor de sănătate (système informatique unique intégré de l'assurance maladie, Roumanie)
SMART	cadre d'autorisation des applications fondées sur des données FHIR (SMART App Launch, SMART on FHIR)
SMD	différence moyenne standardisée (angl. *standardized mean difference*)
SNOMED CT	terminologie clinique systématisée de la médecine (angl. *Systematized Nomenclature of Medicine Clinical Terms*)
SNR	rapport signal/bruit (angl. *signal-to-noise ratio*)
SOP	paire service-objet dans DICOM (angl. *Service-Object Pair*)
SpO2	saturation périphérique en oxygène, mesurée par oxymétrie de pouls
SQL	langage de requête structuré (angl. *Structured Query Language*)
SRIM	Societatea Română de Informatică Medicală (Société roumaine d'informatique médicale)
STARD-AI	lignes directrices pour la publication des études d'exactitude diagnostique des tests fondés sur l'IA
SWF	profil IHE pour le flux de travail programmé en radiologie (angl. *Scheduled Workflow*)
TDM, CT	tomodensitométrie (angl. *computed tomography*)
TEP, PET	tomographie par émission de positons (angl. *positron emission tomography*)
TIC	technologies de l'information et de la communication
TLS	protocole de sécurité de la couche transport (angl. *Transport Layer Security*)
TP	travaux pratiques
TRIPOD+AI	lignes directrices pour la publication du développement et de la validation des modèles prédictifs, y compris ceux construits par apprentissage automatique
TVA	taxe sur la valeur ajoutée
UE	Union européenne
UH	unités Hounsfield
UID	identifiant unique au niveau mondial d'un objet DICOM (angl. *unique identifier*)
UMLS	système unifié de langage médical (angl. *Unified Medical Language System*)
USMLE	examen américain d'autorisation d'exercice de la médecine (angl. *United States Medical Licensing Examination*)
UTF-8	forme de codage Unicode sur 1-4 octets (angl. *Unicode Transformation Format, 8 bits*)
VFC, HRV	variabilité de la fréquence cardiaque (angl. *heart rate variability*)
VN	vrai négatif
VNA	archive neutre vis-à-vis des fournisseurs (angl. *vendor neutral archive*)
VP	vrai positif
VPN	valeur prédictive négative
VPP	valeur prédictive positive
WFDB	format et programmes pour les bases de données de formes d'onde (angl. *Waveform Database*)
WHA	Assemblée mondiale de la santé (angl. *World Health Assembly*)
WSI	numérisation de lames entières en pathologie numérique (angl. *whole slide imaging*)
XAI	IA explicable (angl. *explainable AI*)
XML	langage de balisage extensible (angl. *eXtensible Markup Language*)

PARTIE I

Partie I. Cours

CHAPITRE 1

Chapitre 1. L'informatique médicale : objet, domaines, compétences

Objectifs

À l'issue de ce chapitre, l'étudiant sera capable :

1. de définir l'informatique biomédicale, l'informatique médicale, l'informatique de la santé et l'informatique clinique et d'expliquer ce qui les distingue.
2. de distinguer, sur un exemple clinique, une donnée d'une information et d'une connaissance.
3. de ranger dans l'ordre chronologique au moins huit jalons de la discipline, des cartes perforées aux grands modèles de langage.
4. de décrire les domaines de l'informatique biomédicale sur le spectre molécules, tissus, patient, population.
5. d'énumérer les domaines de compétence numérique exigés du médecin et d'estimer son propre niveau.
6. de construire une requête PubMed avec des termes MeSH et des opérateurs booléens et de vérifier une référence proposée par un outil d'intelligence artificielle.

1.1 Définitions et objet

1.1.1 Du "calcul médical" à l'informatique biomédicale

Le nom de la discipline a changé plusieurs fois au cours des soixante dernières années, et chaque changement dit quelque chose de la manière dont on a compris son objet. Dans les années 1970, on parlait, dans le monde anglophone, de *medical computer science*, c'est-à-dire de la subdivision de l'informatique qui applique les méthodes générales du domaine à des problèmes médicaux. Dans les mêmes années 1970, sur le modèle du français *informatique*, est apparu le terme *medical informatics*, qui a gagné du terrain surtout en Europe dans les années 1980 [7]. Le nouveau terme déplaçait l'accent de l'ordinateur vers l'information. Il englobait la statistique médicale, la tenue des registres et l'étude de la nature de l'information médicale, et pas seulement la programmation [7].

L'adjectif "médical" a ensuite été critiqué comme trop lié au médecin et à la maladie. Sont ainsi apparus *health informatics* et, avec le projet du génome humain des années 1990, *bioinformatics*. Pour englober ces deux orientations, une partie de la communauté a adopté la dénomination d'informatique biomédicale (*biomedical informatics*, BMI). Un signe de ce changement est la revue *Computers and Biomedical Research*, parue en 1967, qui est devenue en 2001 le *Journal of Biomedical Informatics* [7].

La définition adoptée par l'Association américaine d'informatique médicale (AMIA) et reprise par le manuel de référence du domaine décrit l'informatique biomédicale comme le domaine interdisciplinaire qui étudie et vise l'utilisation efficace des données, des informations et des

connaissances biomédicales pour la recherche scientifique, la résolution de problèmes et la prise de décision, dans le but d'améliorer la santé humaine [7]. Cette définition comporte quatre corollaires. Le premier fixe l'étendue du domaine, des molécules aux individus et aux populations, des systèmes biologiques aux systèmes sociaux. Le deuxième concerne la théorie et les méthodes de production, de stockage, de récupération, d'utilisation et de partage des données et des connaissances. Le troisième montre que la discipline s'appuie sur l'informatique, les télécommunications et les sciences de l'information, et qu'elle y contribue. Le quatrième rappelle que les êtres humains sont les utilisateurs finaux, c'est pourquoi la conception et l'évaluation des solutions font aussi appel aux sciences sociales et comportementales [7].

1.1.2 Termes apparentés

Les termes du tableau 1.1 apparaissent fréquemment dans la littérature et ne sont pas synonymes. Shortliffe et Chiang proposent une convention simple : l'informatique biomédicale est la discipline de base, et les autres dénominations désignent ses domaines d'application [7]. Selon cette convention, le terme "informatique médicale" désigne la recherche appliquée et la pratique centrées sur la maladie et sur le rôle du médecin, tandis que le terme "informatique de la santé" recouvre l'informatique clinique et celle de la santé publique prises ensemble.

Tableau 1.1. Termes utilisés pour désigner la discipline et ses domaines. Source : [7-9].

Terme	Terme en anglais	Ce qu'il recouvre
Informatique biomédicale	biomedical informatics	discipline de base : méthodes, théories, outils
Informatique médicale	medical informatics	applications centrées sur la maladie et sur l'activité du médecin
Informatique de la santé	health informatics	informatique clinique et informatique de santé publique, ensemble
Informatique clinique	clinical informatics	applications pour les soins au patient individuel
Bio-informatique	bioinformatics	information biologique, surtout au niveau moléculaire
Informatique de santé publique	public health informatics	applications pour les populations, la prévention, la surveillance

Dans les facultés de médecine de Roumanie, la discipline enseignée aux étudiants porte le nom d'informatique médicale. Le premier cours, introduit à titre optionnel en 1984, est devenu obligatoire pour tous les étudiants en médecine en 1990 [10]. Dans cet ouvrage, le terme "informatique médicale" désigne la discipline telle que la rencontre l'étudiant en médecine, et le terme "informatique biomédicale" apparaît lorsqu'il est question de la science de base.

1.1.3 Données, information, connaissance

Une donnée clinique (*datum*) est une observation unique relative à un patient : une température, une numération des érythrocytes, un antécédent de rubéole. Shortliffe et Chiang montrent qu'une donnée se définit par cinq éléments : le patient, le paramètre observé, la valeur, le moment de l'observation et la méthode par laquelle elle a été obtenue [11]. L'information est constituée de données analysées, organisées de façon à prendre sens. La connaissance résulte de l'analyse, formelle ou informelle, de l'information et comprend aussi bien les résultats des études que des heuristiques et des modèles [11].

Un exemple aide à y voir clair. Une tension artérielle de 180/110 mmHg mesurée chez un patient est une donnée. Lorsque de telles données sont réunies pour de nombreux patients, l'information apparaît, et son analyse peut montrer que les patients dont la tension est élevée font plus souvent un infarctus du myocarde, ce qui constitue une connaissance [11]. La frontière dépend du contexte. "Le patient est hypertendu" est une interprétation de plusieurs mesures, mais devient une simple donnée d'entrée pour un programme d'aide à la décision. Une base de données (*database*) conserve des observations individuelles, sans analyse. Une base de connaissances (*knowledge base*) contient des faits, des heuristiques et des modèles utilisés pour résoudre des problèmes [11]. Le chapitre 2 reprend ces notions avec la mesure de l'information.

1.1.4 Discipline scientifique et pratique professionnelle

L'informatique biomédicale peut être considérée comme une science biomédicale fondamentale, dotée de nombreux domaines d'application. C'est une science expérimentale : elle formule des questions, conçoit des expériences, analyse les résultats et utilise ce qu'elle a appris pour de nouvelles expériences [7]. La recherche fondamentale et la recherche appliquée se prolongent l'une l'autre. Les problèmes naissent de la pratique médicale, et les solutions généralisables y retournent. La discipline s'appuie sur plusieurs sciences constitutives, et pas seulement sur l'informatique : les sciences de la décision, la statistique, les sciences cognitives, les sciences de l'information et les sciences de gestion [7].

La relation avec le génie biomédical mérite une précision. Les départements de génie biomédical sont apparus à la fin des années 1960 et se sont occupés principalement d'instrumentation, de monitorage et de techniques d'imagerie. L'informatique biomédicale, plus jeune, place au centre l'information et la connaissance biomédicales ainsi que leur gestion à l'aide de l'ordinateur [7]. Les deux disciplines se rencontrent souvent, par exemple dans le traitement des signaux et des images (chapitres 6 et 7).

L'informatique est aussi une profession. Aux États-Unis, l'informatique clinique a été approuvée en 2013 comme surspécialité médicale, avec un examen de certification [7]. Pour le médecin qui ne deviendra pas spécialiste en informatique, l'enjeu est autre : utiliser correctement les outils numériques, en reconnaître les limites et participer à leur choix et à leur évaluation. La même source rappelle que le passage d'une innovation dans la pratique courante peut prendre en moyenne 17 ans [7]. C'est précisément pour cette raison qu'il importe que le médecin sache juger les preuves.

1.2 Repères historiques : des premiers ordinateurs à l'hôpital à l'intelligence artificielle générative

L'histoire de l'informatique médicale est aussi celle des ordinateurs. La première application du calcul automatique pertinente pour la médecine a été le système à cartes perforées de Herman Hollerith, utilisé lors du recensement américain de 1890 puis adapté à des enquêtes épidémiologiques [7]. Les ordinateurs électroniques à programme enregistré sont apparus à la fin des années 1940, et les ordinateurs à usage général ont été commercialisés à partir du milieu des années 1950 [7]. C'est aussi à cette époque que l'intelligence artificielle (IA) a pris forme. La conférence de Dartmouth de 1956 a marqué sa constitution en discipline formelle [12].

Dans les années 1960, plusieurs centres américains ont construit les premiers systèmes informatiques hospitaliers. Au Massachusetts General Hospital, Barnett et ses collaborateurs ont développé des applications pendant trois décennies, et le langage MUMPS, créé spécialement

pour des applications médicales, a été pendant quelques années le plus utilisé pour le traitement des dossiers médicaux [7]. La documentation scientifique s'est informatisée en parallèle. L'Index Medicus imprimé est resté la principale source de recherche bibliographique jusqu'en 1971, année où la Bibliothèque nationale de médecine des États-Unis a mis à disposition la version électronique MEDLARS, avec des notices cataloguées depuis 1966, suivie de MEDLINE (*MEDLARS Online*) [13].

Les années 1970 ont apporté les premiers systèmes d'aide à la décision. Le groupe dirigé par de Dombal a utilisé en 1972 un modèle bayésien pour le diagnostic de la douleur abdominale aiguë. MYCIN, un système à base de règles consacré aux infections, n'a jamais été utilisé en clinique, mais a servi de prototype à de nombreux systèmes ultérieurs [14]. Les mini-ordinateurs du début des années 1970 et le microprocesseur, suivi de l'ordinateur personnel, à la fin des années 1970 et au début des années 1980, ont amené le calcul dans chaque service, puis sur chaque bureau [7].

À mesure que les systèmes se sont multipliés, le besoin de les faire communiquer s'est accru. L'organisation HL7 a été fondée en 1987, et dans les années 1990 sont apparus le comité européen de normalisation pour l'informatique de santé et le comité technique ISO/TC 215 [15]. Le standard HL7 FHIR, introduit en 2011, a simplifié l'échange de données grâce à des ressources décrites en XML ou en JSON [15]. En 2009, la loi américaine HITECH a offert des incitations financières aux hôpitaux et aux médecins qui adoptaient des dossiers électroniques certifiés et les utilisaient effectivement [7].

L'apprentissage profond a dépassé, au début des années 2000, bon nombre des limites des premiers modèles d'IA [16]. En 2021, AlphaFold a prédit la structure tridimensionnelle des protéines avec une exactitude comparable, dans la majorité des cas testés, à celle des structures déterminées expérimentalement. Le problème était ouvert depuis plus de 50 ans [17]. Depuis la fin de 2022, ChatGPT a rendu les grands modèles de langage (*large language models*, LLM) accessibles à tous [18]. En 2023, ChatGPT a obtenu aux trois étapes de l'examen américain d'autorisation d'exercice de la médecine (USMLE) des scores égaux ou proches du seuil de réussite [19]. La même année, le modèle Flan-PaLM a atteint une exactitude de 67,6 % sur des questions de type USMLE, et sa variante adaptée au domaine médical, Med-PaLM, a donné des résultats encourageants, mais est restée inférieure aux cliniciens lors de l'évaluation humaine [20].

La pandémie de COVID-19 a accéléré la télémédecine, l'échange de données et l'utilisation de l'IA. Une analyse de 2023 conclut que la santé numérique a gagné en maturité pendant la pandémie, mais que des problèmes sociotechniques, d'évaluation et de confiance subsistent [21]. L'Union européenne a répondu par la réglementation : le règlement sur l'intelligence artificielle a été adopté en 2024 [22], et le règlement relatif à l'espace européen des données de santé en 2025 [23]. La figure 1.1 résume ces repères.

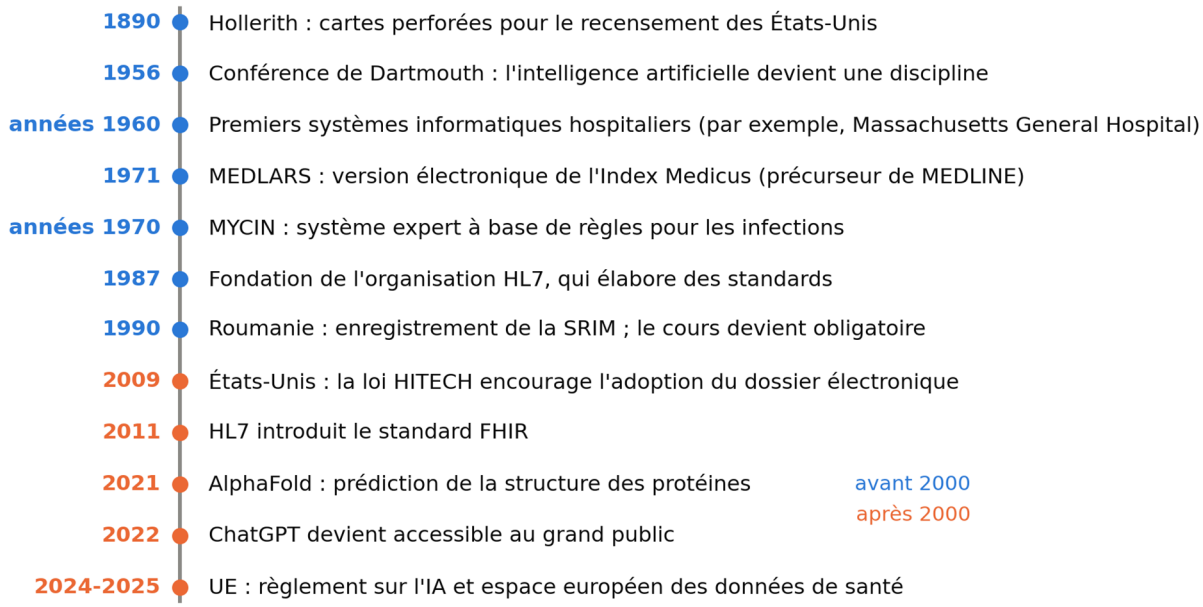


Figure 1.1. Repères de l'histoire de l'informatique médicale. L'axe n'est pas proportionnel au temps. Sources : [7,10,12-15,17,18,22,23].

1.3 Les domaines de l'informatique biomédicale : des molécules aux populations

Les mêmes méthodes (représentation des données, standards, raisonnement, évaluation) s'appliquent à différents niveaux d'organisation du vivant. Shortliffe et Chiang décrivent un spectre à quatre degrés : processus moléculaires et cellulaires, tissus et organes, individu (patient), populations et société [7]. À chaque degré correspond un domaine d'application, comme le montre la figure 1.2. Les frontières sont floues. La pharmacogénomique, par exemple, relie des bases de données génotypiques et phénotypiques et se situe à l'intersection de la bio-informatique et de l'informatique clinique [7].

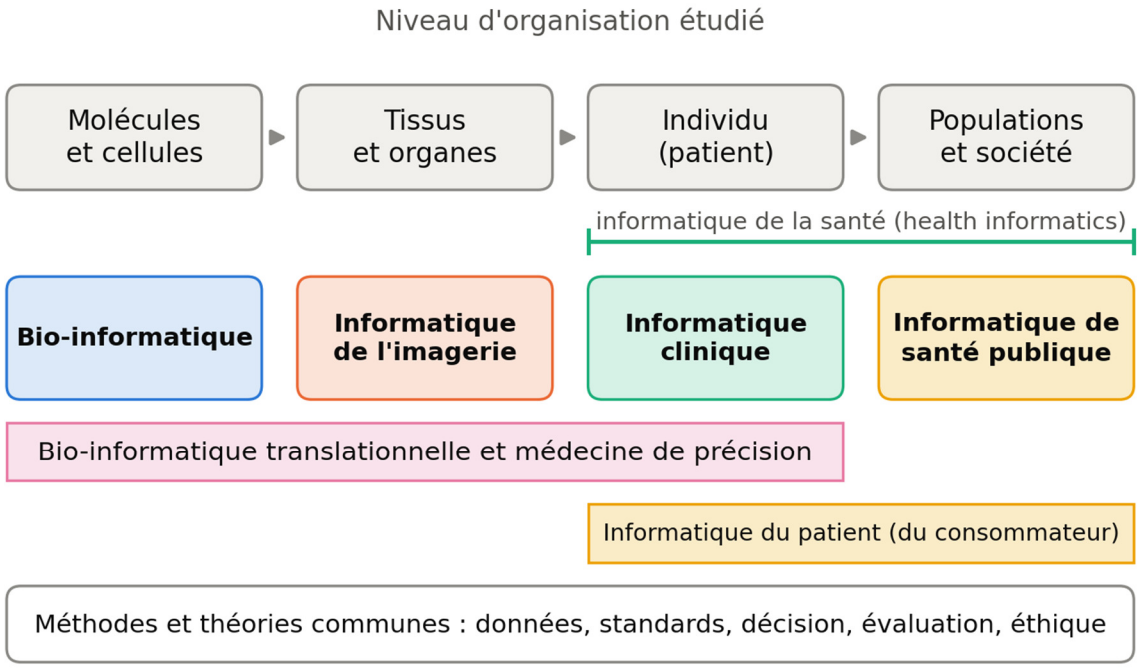


Figure 1.2. Les domaines de l'informatique biomédicale sur le spectre allant des molécules aux populations. Adapté de [7].

1.3.1 La bio-informatique

La bio-informatique étudie la manière dont l'information est représentée et analysée dans les systèmes biologiques, surtout au niveau moléculaire. Si l'informatique clinique gère l'information liée aux soins des patients, la bio-informatique s'occupe de l'information issue des sciences biologiques fondamentales [8]. Les données de séquençage, d'expression génique ou de métabolomique sont aujourd'hui produites à un rythme que seules les méthodes computationnelles peuvent suivre. AlphaFold est un exemple de résultat bio-informatique ayant un impact direct sur la recherche médicale [17].

1.3.2 L'informatique de l'imagerie

Les images médicales sont obtenues directement au format numérique ou converties dans ce format. L'informatique de l'imagerie s'occupe de l'acquisition des images, de la représentation de leur contenu, de leur stockage et de leur gestion, de leur traitement et de leur interprétation, éventuellement assistée par ordinateur [24]. Le nombre d'images produites lors d'un seul examen peut atteindre plusieurs centaines ou plusieurs milliers [24]. Le chapitre 7 présente en détail le standard DICOM et le traitement des images.

1.3.3 L'informatique clinique

L'informatique clinique couvre les applications orientées vers le patient individuel, en médecine, en soins infirmiers ou en médecine dentaire [7]. Le dossier de santé électronique (DSE, en anglais *electronic health record*, EHR), la prescription électronique, les systèmes d'aide à la décision clinique et la télémédecine relèvent de ce domaine. Ils font l'objet des chapitres 8-12.

1.3.4 L'informatique de santé publique

L'informatique de santé publique applique de façon systématique les méthodes et les outils de l'informatique pour soutenir les objectifs de santé publique, quel que soit le cadre dans lequel on travaille [9]. La santé publique remplit trois fonctions essentielles : l'évaluation de l'état de santé de la population, l'élaboration des politiques et la garantie de l'accès aux services [9]. Chacune exige d'autres données et d'autres systèmes. La pandémie a montré à quel point les dossiers électroniques et l'informatique de santé publique sont étroitement liés dans la gestion d'une épidémie [7]. Dans le même contexte est discuté le concept de système de santé apprenant (*learning health system*), défini comme un système qui utilise en continu une infrastructure fondée sur les données pour les politiques, la santé publique et la personnalisation des soins [25].

1.3.5 L'informatique du patient (du consommateur)

Les soins centrés sur le patient sont devenus une composante essentielle de la médecine moderne, et l'informatique du patient (*personal* ou *consumer health informatics*) étudie les outils par lesquels la personne gère sa propre santé : portails patients, dossiers personnels de santé, applications mobiles, capteurs à domicile [26]. Relèvent aussi de ce domaine les problèmes de littératie en santé et de fracture numérique [26].

1.3.6 L'informatique translationnelle

La bio-informatique translationnelle applique les méthodes de la bio-informatique à la santé et à la maladie humaines. Elle aide les découvertes de laboratoire à parvenir au lit du malade et, en définitive, à la santé de la population, par la médecine de précision [27]. À la différence de la bio-informatique "classique", elle travaille principalement sur des données humaines, c'est pourquoi

elle rencontre des problèmes de réglementation et de confidentialité qui ne se posent pas pour les données provenant d'organismes modèles [27].

La protection des données traverse tous ces domaines. Une analyse bibliométrique a recensé 7 165 articles sur la confidentialité des données en informatique médicale, publiés entre 1975 et 2023, avec un nombre annuel en hausse. Les thèmes les plus étudiés ont été la santé mobile et la télémédecine, la sécurité et l'authentification, ainsi que l'informatique en nuage (*cloud computing*) [28]. Le chapitre 5 traite de ces aspects.

1.4 Les compétences numériques du médecin

Que doit savoir un médecin en matière d'informatique ? La réponse a changé à mesure que les outils ont pénétré chaque étape de l'acte médical. Plusieurs organisations ont publié des référentiels de compétences, qui se recoupent en grande partie.

1.4.1 Les recommandations de l'IMIA

L'Association internationale d'informatique médicale (IMIA) a publié des recommandations pour l'enseignement de l'informatique biomédicale et de santé en 2000 et en 2010, et leur deuxième révision a paru en 2023 [29]. Cette révision a été élaborée par un groupe de travail réunissant des représentants de toutes les régions de l'IMIA, sur la base d'une étude Delphi internationale. Elle décrit les besoins éducatifs de trois rôles (utilisateur, généraliste et spécialiste) dans six domaines : les principes de base de la discipline, les sciences de la santé et les services de santé, l'informatique et les sciences des données et de l'information, les sciences sociales et comportementales, les sciences de gestion et, enfin, la spécialisation en informatique [29]. Pour l'étudiant en médecine, le rôle pertinent est avant tout celui d'utilisateur compétent.

1.4.2 Les référentiels consacrés à l'enseignement médical

Le référentiel DECODE (*Digital Health Competencies in Medical Education*), publié en 2025, est issu d'une étude Delphi menée auprès de 211 experts de 79 pays et territoires. Il regroupe les compétences en quatre domaines : le professionnalisme en santé numérique, la santé numérique du patient et de la population, les systèmes d'information de santé et la science des données de santé. Les 19 compétences comportent 33 acquis d'apprentissage obligatoires et 145 optionnels [30]. Une étude antérieure, fondée sur une revue de la littérature et sur un panel Delphi de 18 experts, a retenu 40 des 65 thèmes de santé numérique comme nécessaires dans le cursus des étudiants en médecine [31].

L'attitude des étudiants est en général favorable. Une méta-analyse de 2026, portant sur 96 études transversales menées dans 37 pays auprès de plus de 45 000 étudiants, a estimé à 76,9 % la proportion d'étudiants ayant une attitude positive envers l'IA et à 76,6 % celle qui soutenait son intégration dans le cursus. Dans le même temps, 39,9 % craignaient que les médecins puissent être remplacés, et la variation entre les centres était très importante [32].

1.4.3 Le cadre européen et la stratégie de l'OMS

Le cadre européen des compétences numériques pour les citoyens, DigComp, a été publié dans sa version 2.2 en 2022 [33]. La version 3.0, publiée le 27 novembre 2025, conserve les cinq domaines et les 21 compétences, mais intègre dans tous les domaines les compétences liées à l'IA [34]. Les cinq domaines sont la recherche, l'évaluation et la gestion de l'information, la communication et la collaboration, la création de contenu, la sécurité (avec le bien-être et l'utilisation responsable) et, enfin, l'identification et la résolution de problèmes [34].

L'Observatoire européen des systèmes et des politiques de santé a analysé en 2025, dans le Policy Brief 71, les compétences numériques des professionnels de santé de l'UE. Les référentiels de compétences analysés mettent l'accent sur la maîtrise de l'information et des données, la communication et la collaboration, la sûreté et la sécurité, la résolution de problèmes et la création de contenu numérique [35]. L'offre de formation est jugée insuffisante, avec des lacunes dans les domaines avancés (robotique, IA, génomique) et une dépendance excessive aux cours en anglais. Les participants d'Italie et de Roumanie aux groupes de discussion ont explicitement signalé que le manque de connaissances en anglais limite l'accès aux cours [35]. Le document présente aussi un exemple de réglementation. La France a adopté un référentiel commun de compétences numériques pour les professions de santé, comportant cinq domaines, et a imposé par la loi son intégration dans la formation de 12 professions de l'enseignement supérieur d'ici 2024-2025 [35].

À l'échelle mondiale, la Stratégie mondiale de l'OMS pour la santé numérique 2020-2025 a été approuvée par l'Assemblée mondiale de la santé en 2020. Parmi les actions proposées aux États membres figure l'intégration des compétences en santé numérique dans les programmes de formation de tous les professionnels de santé [36]. L'Assemblée mondiale de la santé a approuvé en mai 2025 la prolongation de la stratégie jusqu'en 2027 [37,38]. À cette date, 129 pays disposaient d'une stratégie nationale de santé numérique [37].

Le tableau 1.2 traduit ces référentiels en tâches concrètes pour la première année et indique où elles sont mises en pratique dans l'ouvrage.

Tableau 1.2. Domaines de compétence numérique pertinents en première année. Source : adapté de [30,34,35].

Domaine (DigComp)	Ce que cela signifie pour l'étudiant en médecine	Où s'exercer
Information et données	recherche PubMed avec le MeSH, évaluation des sources, gestion des références	chap. 1, 2 ; TP 1, TP 2
Communication et collaboration	partage contrôlé, travail en équipe sur des documents, téléconsultation	TP 1, chap. 9
Création de contenu	documents, tableaux, graphiques, citation correcte	TP 3, TP 5
Sécurité et utilisation responsable	mots de passe, validation en deux étapes, RGPD, éthique de l'IA	TP 1, chap. 5, 11
Résolution de problèmes	choix de l'outil approprié, vérification des résultats de l'IA	chap. 10, 13

1.5 La documentation scientifique

1.5.1 PubMed, MEDLINE et MeSH

PubMed est le moteur de recherche gratuit de la Bibliothèque nationale de médecine des États-Unis et comprend plus de 40 millions de citations issues de MEDLINE, de revues des sciences de la vie et de livres en ligne [39]. Le 27 septembre 2026, une requête portant sur l'ensemble de la base a renvoyé 41 213 139 enregistrements [39]. MEDLINE est la composante indexée à l'aide du vocabulaire contrôlé MeSH (*Medical Subject Headings*). Dans l'édition décrite par le manuel cité, le MeSH comprenait plus de 28 000 descripteurs et plus de 90 000 synonymes (*entry terms*), organisés hiérarchiquement en 16 arborescences, telles que Maladies, Organismes ou Produits chimiques et médicaments [13]. La hiérarchie permet de faire "exploser" un terme, c'est-à-dire d'inclure automatiquement les termes plus spécifiques. Les qualificatifs (*subheadings*)

restreignent le sens, par exemple au diagnostic ou au traitement, et le type de publication permet de sélectionner les articles de synthèse, les essais randomisés ou les méta-analyses [13].

L'indexation s'est progressivement automatisée. L'outil d'indexation automatique de la NLM a évolué d'un système à base de règles, en 2002, vers MTIX, fondé sur des réseaux de neurones, en 2024. Une revue de 2026 montre qu'environ un tiers des enregistrements nécessitent encore une correction humaine, c'est pourquoi celui qui effectue la recherche doit rester attentif aux ambiguïtés algorithmiques [40].

La figure 1.3 utilise PubMed comme source de données. Le nombre annuel d'enregistrements indexés avec le descripteur "Medical Informatics" (qui inclut, par explosion, tous les termes subordonnés) a augmenté jusqu'en 2019, a ensuite diminué jusqu'à un minimum en 2022-2023, puis s'est partiellement redressé en 2024-2025. Le nombre d'enregistrements indexés avec "Artificial Intelligence" ou "Machine Learning" a été multiplié par environ 14 entre 2005 et 2025, passant de 3 579 à 51 707. La baisse de la série consacrée à l'informatique médicale ne doit pas être interprétée automatiquement comme une baisse de l'intérêt. Elle peut aussi refléter des changements d'indexation, et les enregistrements récents peuvent recevoir leurs descripteurs MeSH avec retard. Il s'agit d'une hypothèse, et non d'un résultat vérifié dans ce chapitre.

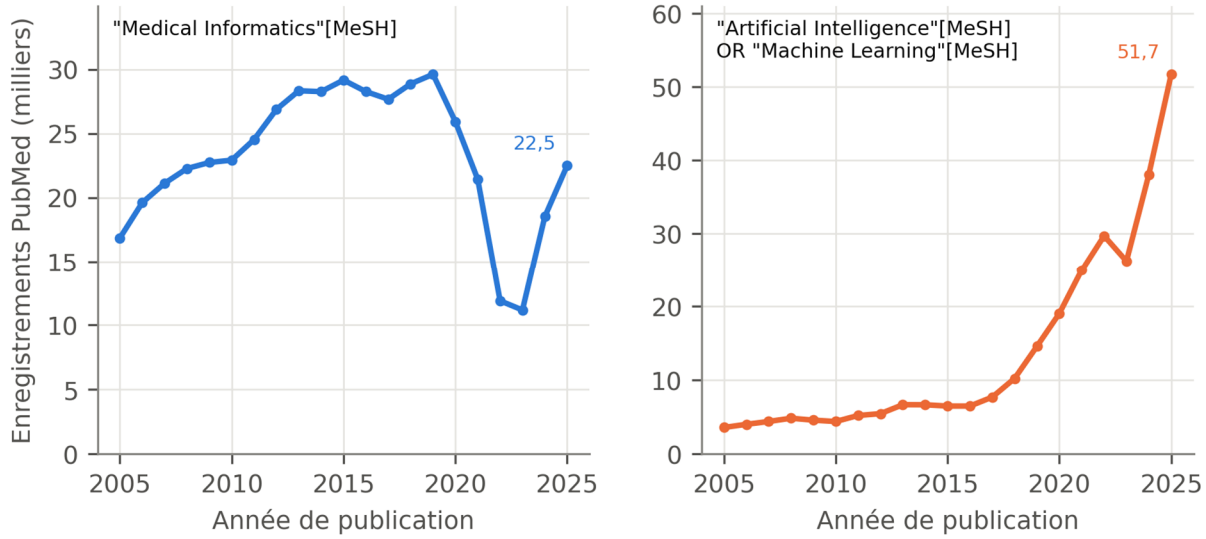


Figure 1.3. Nombre annuel d'enregistrements PubMed indexés avec le descripteur "Medical Informatics" (à gauche) et avec "Artificial Intelligence" ou "Machine Learning" (à droite), 2005-2025. Les axes verticaux ont des échelles différentes. Source : PubMed (NCBI), requête effectuée le 27 septembre 2026 via E-utilities ; calculs de l'auteur [39]. Les données se trouvent dans note/date_pubmed_C01.csv.

1.5.2 Opérateurs booléens, champs et filtres

La recherche dans PubMed combine des ensembles de résultats au moyen d'opérateurs booléens. AND restreint les résultats aux documents qui contiennent tous les concepts. OR réunit les différentes manières d'exprimer un même concept, et NOT retire un ensemble d'un autre [13]. Les opérateurs s'écrivent en majuscules, la recherche est traitée de gauche à droite, et les parenthèses regroupent les concepts [41]. Les balises de champ précisent où chercher : [mh] pour les termes MeSH, [majr] pour le sujet principal, [tiab] pour le titre et le résumé, [pt] pour le type de publication, [dp] pour la date de publication. L'astérisque tronque un terme (au moins quatre caractères doivent le précéder), et les guillemets permettent de rechercher une expression exacte [41].

Lorsque l'utilisateur saisit quelques mots dans la zone de recherche, PubMed tente de les faire correspondre successivement à des termes MeSH, à des titres de revues, à des expressions fréquentes et à des auteurs. Ce qui reste sans correspondance est recherché en texte libre [13]. Les filtres situés à gauche de la page limitent les résultats selon le type d'article, la date de publication, l'espèce, la langue ou l'âge, et la fonction Clinical Queries applique des stratégies validées pour les questions de thérapeutique, de diagnostic, d'effets nocifs et de pronostic [13,41]. Le tri par défaut, Best Match, repose sur un algorithme de pertinence [41]. Le TP 2 met en pratique ces étapes sur une question clinique.

1.5.3 La pyramide des preuves et les sources secondaires

Tous les articles n'ont pas le même poids. La hiérarchie classique des preuves part des études sur l'animal, de la recherche translationnelle et de l'opinion d'experts, s'élève par les rapports et les séries de cas, puis par les études observationnelles analytiques, atteint les essais contrôlés randomisés et place à son sommet les revues systématiques et les méta-analyses [42]. Cette hiérarchie aide à établir des priorités de lecture, mais ne remplace pas l'évaluation critique de chaque article [42]. La pratique fondée sur les preuves s'est déplacée, au fil du temps, de l'évaluation critique des études individuelles vers l'utilisation de preuves déjà synthétisées, c'est-à-dire les revues systématiques, les recommandations et les synthèses intégrées aux systèmes cliniques [13]. Pour l'étudiant de première année, la règle pratique est simple : on cherche d'abord une synthèse récente et une recommandation, puis les études primaires.

1.5.4 Les outils de recherche assistés par l'IA et leurs limites

Les chatbots (agents conversationnels) généralistes produisent des textes fluides, avec des références qui paraissent crédibles. Plusieurs études ont vérifié ces références une à une. Dans une étude de 2023, ChatGPT-3.5 a généré 30 textes médicaux courts comportant 115 références. Parmi celles-ci, 47 % étaient inventées, 46 % existaient mais contenaient des erreurs, et seulement 7 % étaient authentiques et exactes. Un PMID erroné était l'erreur la plus fréquente [43]. Lorsqu'on leur a demandé de refaire 11 revues systématiques, GPT-3.5, GPT-4 et Bard ont présenté des taux d'hallucination de 39,6 %, 28,6 % et 91,4 % respectivement, avec une précision de la sélection des articles d'au plus 13,4 % [44]. Dans une étude interdisciplinaire portant sur ChatGPT (GPT-3.5), environ trois quarts des citations existaient, mais le DOI n'était correct que dans 32,7 % des cas en sciences de la nature et dans 8,5 % des cas en sciences humaines [18].

Les solutions techniques réduisent le problème, sans l'éliminer. Un modèle spécialisé qui récupère d'abord des passages d'articles réels et ne rédige qu'ensuite la réponse (*retrieval-augmented generation*, RAG), OpenScholar, a obtenu une exactitude des citations comparable à celle des experts, tandis que GPT-4o a inventé des citations dans 78-90 % des cas [45]. Dans une étude randomisée menée auprès de 30 urologues, un chatbot avec RAG et références dans le texte a donné des recommandations correctes dans 73 % des cas, contre 50 % pour ChatGPT, et 28 % des citations de ChatGPT étaient inexistantes ou obsolètes [46]. Une revue systématique de 519 études d'évaluation des LLM en santé a constaté que seulement 5 % d'entre elles avaient utilisé des données réelles issues des soins aux patients [47].

La conclusion pratique pour l'étudiant est la suivante : un outil d'IA peut aider à formuler la question ou à trouver des synonymes, mais chaque référence doit être vérifiée dans PubMed ou par son DOI avant d'être citée. Le même principe a d'ailleurs été appliqué lors de la rédaction de cet ouvrage.

Tableau 1.3. Exactitude des références produites par des modèles de langage, dans des études publiées en 2023-2026. Source : [43-46].

Étude	Tâche	Résultat principal
Bhattacharyya 2023	115 références dans des textes médicaux (GPT-3.5)	47 % inventées, 7 % correctes
Chelli 2024	reproduction de 11 revues systématiques	hallucination 28,6-91,4 %
Asai 2026	synthèse de la littérature, sans et avec RAG	GPT-4o : 78-90 % de citations inventées
Carl 2026	cas d'uro-oncologie, ECR avec 30 urologues	28 % de citations ChatGPT inexistantes ou obsolètes

1.6 L'informatique médicale en Roumanie et à Timișoara

Le premier événement roumain consacré à l'informatique médicale a eu lieu en 1977. La Société roumaine d'informatique médicale (Societatea Română de Informatică Medicală, SRIM) n'a pu être officiellement enregistrée qu'en 1990, après la chute du régime communiste [10]. La Roumanie a été membre correspondant de l'IMIA à partir de 1986 et membre de plein droit à partir de 1994, année où la SRIM est aussi devenue membre de la Fédération européenne d'informatique médicale (EFMI) [48]. Deux représentants de la SRIM ont ensuite dirigé l'EFMI : George Mihalaș (2006-2008) et Lăcrămioara Stoicu-Tivadar (2018-2020). La SRIM a organisé des conférences thématiques de l'EFMI à Bucarest, en 2001, et à Timișoara [10]. Le siège de la société se trouve à Timișoara [49]. En mai 2026, la SRIM a organisé à Bucarest la deuxième édition de la conférence jRoMedINF, consacrée au cycle de vie des données médicales [49].

Timișoara a également accueilli la conférence thématique de l'EFMI du 27 au 29 novembre 2024, organisée par la SRIM à l'Université Politehnica de Timișoara, avec l'Université de l'Ouest et l'Université de médecine et de pharmacie "Victor Babeș" comme coorganisatrices. Les résumés étendus des posters ont paru dans un supplément de la revue *Applied Medical Informatics* [50]. La revue, éditée à l'Université de médecine et de pharmacie "Iuliu Hațieganu" de Cluj-Napoca, est en libre accès et ne perçoit pas de frais de publication [51]. L'EFMI mentionne aussi la collaboration de la SRIM avec la Société roumaine d'informatique médicale appliquée [10].

À Timișoara, la discipline a sa propre tradition didactique. Le manuel d'informatique médicale et de biostatistique de Mihalaș et Lungeanu [1] et l'ouvrage récent de Focșa sur l'informatique médicale, des données à l'intelligence artificielle [2], sont des lectures de base pour ce cours.

La normalisation est la troisième composante de la vie professionnelle. En Roumanie, les normes d'informatique médicale sont suivies par le comité technique ASRO/CT 319 "Informatică medicală" (informatique médicale), correspondant national des comités ISO/TC 215 et CEN/TC 251 [52]. L'auteur de l'ouvrage participe aux travaux du comité technique de l'ASRO pour l'informatique médicale.

La situation en Roumanie en 2026

- Le cours d'informatique médicale est obligatoire pour tous les étudiants en médecine depuis 1990 [10]. À l'UMF "Victor Babeș" de Timișoara, il est enseigné en première année, selon le programme de 2025.
- La SRIM, dont le siège est à Timișoara, est membre de l'IMIA et de l'EFMI depuis 1994 [48,49].
- Les normes ISO et CEN d'informatique médicale sont suivies par l'ASRO/CT 319 [52].
- Les documents européens signalent la barrière de la langue anglaise dans l'accès des professionnels roumains à la formation numérique [35].

1.7 Comment utiliser cet ouvrage

L'ouvrage suit le programme analytique de 2025 de la discipline : 14 cours et 14 séances de travaux pratiques, soit une paire par semaine. Chaque chapitre de cours se termine par des questions à choix multiples, et chaque séance de travaux pratiques décrit ce qui doit être rendu pour l'examen pratique. Le tableau 1.4 présente cette correspondance.

Tableau 1.4. Correspondance entre les cours, les travaux pratiques et les chapitres de l'ouvrage (programme de 2025).

Sem.	Cours	Travaux pratiques	Chapitre
1	Objet de l'informatique médicale ; documentation	Matériel et logiciels ; Google Workspace	Chap. 1
2	Théorie de l'information ; données et connaissances	Sources web, évaluation de la qualité	Chap. 2
3	Fichiers, SGBD, portabilité, NoSQL	Google Forms, Epi Info	Chap. 3
4	CIM-10, ATC, DRG en Roumanie	Codage CIM en ligne	Chap. 4
5	Éthique, RGPD, sécurité des données	Citation, plagiat, essai	Chap. 5
6	Biosignaux	ECG/EEG, logiciels libres	Chap. 6
7	Images biomédicales	ImageJ/Fiji	Chap. 7
8	eHealth I : le dossier électronique	OpenEMR, version de démonstration	Chap. 8
9	eHealth II : télémédecine	Téléconsultation simulée	Chap. 9
10	mHealth, interopérabilité	Consultation au moyen d'un LLM	Chap. 10
11	Système d'information ; signature électronique	Sécurité, signature	Chap. 11
12	Système d'information médical, SIUI, SIH	Cabinets, polycliniques, hôpitaux	Chap. 12
13	L'IA en médecine	Projets et débats sur l'IA	Chap. 13
14	Révision	Examen pratique	Chap. 14

L'ouvrage cite les manuels de la discipline, mais ne les remplace pas [1,2]. L'étudiant les utilise en parallèle, surtout pour la biostatistique et pour les thèmes qui y sont traités en détail. Les méthodes de recherche avancées, comme les jumeaux numériques, l'apprentissage fédéré ou la présentation des modèles prédictifs selon TRIPOD+AI, sont traitées dans la monographie doctorale de l'auteur.

Pièges fréquents

- Croire que l'informatique médicale consiste à "savoir se servir d'un ordinateur". Son objet est l'information et la décision, et l'ordinateur n'est qu'un instrument.
- Employer "bio-informatique" comme synonyme de l'ensemble de la discipline. La bio-informatique n'est que le domaine qui travaille au niveau moléculaire et cellulaire.
- Chercher dans PubMed comme dans Google, avec une longue phrase, sans MeSH et sans opérateurs. Le résultat est soit trop large, soit incomplet.
- Écrire "and" au lieu de AND ou oublier les parenthèses lorsqu'on combine OR et AND.
- Citer une référence proposée par un chatbot sans l'avoir retrouvée dans PubMed ou par son DOI. Les études montrent des taux élevés de références inventées.
- Confondre la hiérarchie des preuves et la qualité : une méta-analyse médiocre ne devient pas bonne parce qu'elle se trouve au sommet de la pyramide.

À retenir

- L'informatique biomédicale étudie l'utilisation efficace des données, des informations et des connaissances pour la recherche, la décision et la santé.
- La donnée est une observation, l'information est un ensemble organisé de données, la connaissance résulte de l'analyse de l'information.
- Les domaines d'application s'étendent des molécules (bio-informatique) aux populations (informatique de santé publique).
- Les référentiels de l'IMIA, DECODE et DigComp exigent du médecin des compétences d'information, de communication, de sécurité et de résolution de problèmes.
- On interroge PubMed efficacement avec le MeSH, des opérateurs booléens écrits en majuscules, des parenthèses et des filtres.
- Toute référence produite par un outil d'IA doit être vérifiée à la source.
- En Roumanie, la discipline est obligatoire depuis 1990, et la SRIM a son siège à Timișoara.

Questions d'autoévaluation

1. (QCS) Selon la définition adoptée par l'AMIA, l'informatique biomédicale est :

a) un synonyme de la bio-informatique
b) la branche de l'ingénierie qui conçoit l'appareillage médical
c) le domaine interdisciplinaire qui étudie l'utilisation efficace des données, des informations et des connaissances biomédicales pour la recherche, la résolution de problèmes et la prise de décision, dans le but d'améliorer la santé
d) la discipline qui s'occupe exclusivement du dossier de santé électronique
e) une surspécialité de la statistique médicale

2. (QCS) L'enregistrement "patient X, tension artérielle 180/110 mmHg, 14 h 30, tensiomètre automatique" représente :

a) une connaissance
b) une donnée clinique
c) une base de connaissances
d) une heuristique
e) une recommandation de pratique

3. (QCS) Dans PubMed, l'opérateur qui ne conserve que les articles contenant les deux concepts est :

a) OR
b) NOT
c) AND
d) l'astérisque (*)
e) les guillemets

4. (QCM) Lesquelles des affirmations suivantes sur les références générées par les grands modèles de langage sont étayées par les études présentées dans ce chapitre ?

a) dans une étude portant sur 115 références produites par ChatGPT-3.5, 47 % étaient inventées
b) tous les modèles testés ont eu des taux d'hallucination inférieurs à 5 %
c) lors de la reproduction de revues systématiques, Bard a eu un taux d'hallucination de 91,4 %
d) un chatbot avec récupération à partir de sources (RAG) et références dans le texte a augmenté la proportion de recommandations correctes dans une étude randomisée
e) le PMID proposé par le modèle est correct dans tous les cas

5. (QCM) Lesquels des éléments suivants sont des domaines de compétence du cadre européen DigComp ?

a) la recherche, l'évaluation et la gestion de l'information
b) la communication et la collaboration
c) la création de contenu
d) la pharmacologie clinique numérique
e) l'identification et la résolution de problèmes

CHAPITRE 2

Chapitre 2. Données, information et connaissance en médecine

Objectifs

À la fin de ce chapitre, l'étudiant pourra :

- classer une donnée médicale selon sa structure (structurée, semi-structurée, non structurée) et selon son échelle de mesure (nominale, ordinale, numérique)
- calculer la quantité d'information d'un résultat et l'entropie d'une variable binaire, à l'aide du logarithme en base 2
- expliquer, sur un exemple chiffré, le lien entre la probabilité pré-test, la probabilité post-test et l'information apportée par un test diagnostique
- estimer l'ordre de grandeur d'un fichier ECG, d'un examen de tomodensitométrie et d'un génome séquencé
- décrire la hiérarchie données-information-connaissance-sagesse et au moins deux de ses limites
- repérer les sources de bruit dans la transmission d'un patient et justifier le recours à un protocole structuré (I-PASS, SBAR).

2.1 Les données médicales

2.1.1 Qu'est-ce qu'une donnée clinique

Au sens informatique, une donnée clinique (*datum*) est une observation unique concernant un patient : une température, une valeur d'hémoglobine, un antécédent de rubéole ou une pression artérielle [11]. Pour qu'une telle observation puisse être traitée par l'ordinateur, elle doit être décrite par cinq éléments : le patient, le paramètre observé, la valeur, le moment de l'observation et la méthode par laquelle elle a été obtenue (déclaration du patient, thermomètre, bandelette urinaire, analyseur de laboratoire) [11]. L'exemple de la pression artérielle montre pourquoi le modèle de données compte. Le médecin lit "120/80" comme une seule valeur, mais pour l'analyse elle représente deux informations distinctes, la pression systolique et la pression diastolique.

Le temps complique beaucoup la gestion des données. En consultation ambulatoire, la date calendaire de la consultation suffit habituellement. En soins intensifs, c'est la minute qui compte, et pour un moniteur de signes vitaux, la fraction de seconde.

2.1.2 Types de données selon la structure

Les données **structurées** ont un format fixe, établi à l'avance : un champ numérique pour la glycémie, un code de diagnostic, une date calendaire. On peut les trier, les compter et les comparer directement. Les données **semi-structurées** comportent des étiquettes, mais le contenu de chaque étiquette peut varier. Un message de laboratoire au format JSON ou XML en est l'exemple type (chapitre 3). Les données **non structurées** sont le texte libre, l'image, l'enregistrement audio. Le

texte narratif (anamnèse, évolution, lettre médicale) représente une part importante de l'information recueillie au cours de la prise en charge des patients [11]. Il est riche en nuances, mais difficile à interroger automatiquement sans méthodes de traitement du langage naturel.

2.1.3 Types de données selon l'échelle de mesure

La statistique utilise une autre classification, fondée sur les opérations qui ont un sens pour les valeurs :

- **nominales** : catégories sans ordre (groupe sanguin, sexe, service d'hospitalisation)
- **ordinales** : catégories ordonnées, mais sans distances égales entre elles (stade NYHA, score de Glasgow par catégories, degré de douleur "légère/modérée/sévère")
- **numériques** : valeurs mesurées sur une échelle dotée d'unités (pression artérielle en mmHg, créatinine en mg/dL), qui peuvent être discrètes (nombre de grossesses) ou continues (poids).

Coder une variable ordinale par les chiffres 1, 2, 3 ne la transforme pas en variable numérique. La moyenne des degrés de douleur, par exemple, a un sens discutable, car "modérée" ne vaut pas nécessairement deux fois "légère".

2.1.4 Formes particulières de données

Outre les valeurs isolées, la médecine produit aussi des données dotées d'une structure propre :

- **séries temporelles** : la même variable mesurée de façon répétée (la glycémie toutes les heures, la température relevée sur la feuille de surveillance)
- **signaux** : variations continues échantillonnées par un appareil (ECG, EEG, oxymétrie de pouls)
- **images** : radiographies, TDM, IRM, lames d'anatomie pathologique numérisées
- **données omiques** : génomique, transcriptomique, protéomique, métabolomique. Le domaine est parti de la génomique pour s'étendre à de nombreuses technologies de mesure à haut débit, et les tendances actuelles sont l'analyse à l'échelle de la cellule unique et l'intégration de plusieurs couches omiques dans une même analyse [53]
- **données générées par le patient** (*patient-generated health data*, PGHD) : données cliniquement pertinentes, recueillies par le patient en dehors du cadre de soins traditionnel. Dans une revue de portée (*scoping review*) qui a inclus 19 études, les données intégrées au dossier électronique étaient biométriques et relatives à l'activité du patient dans 57,9 % des études, et leur intégration dans le dossier de santé électronique (DSE, *electronic health record*, EHR) n'en était qu'à ses débuts [54].

Tableau 2.1. Types de données médicales, avec des exemples et leur forme de stockage habituelle. Source : synthèse d'après [11,53,54].

Type	Exemple clinique	Structure	Stockage typique
Valeur numérique	Potassium sérique 4,1 mmol/L	Structurée	Champ numérique avec unité
Code	Diagnostic codé selon la CIM	Structurée	Champ texte contrôlé
Texte narratif	Anamnèse, compte rendu de sortie	Non structurée	Texte libre
Signal	ECG à 12 dérivations	Série d'échantillons	Fichier binaire
Image	Examen TDM	Matrice de pixels	Fichiers DICOM
Données omiques	Séquence de génome	Chaîne de nucléotides	FASTQ, BAM, VCF
Données générées par le patient	Nombre quotidien de pas, glycémie à domicile	Variable	Application mobile, dispositif portable

2.1.5 La qualité des données

Une donnée erronée conduit à une décision erronée, même si le programme qui la traite est irréprochable. Les dimensions classiques de la qualité sont au nombre de quatre :

- **l'exactitude** (la valeur correspond à la réalité : le poids enregistré est le poids réel)
- **la complétude** (les valeurs nécessaires sont présentes)
- **l'actualité** (la valeur est suffisamment récente pour la décision à prendre)
- **la cohérence** (les valeurs ne se contredisent ni entre elles ni d'un système à l'autre : un patient de sexe masculin ne peut pas avoir de grossesse en cours).

Une revue systématique de 103 travaux publiés entre 2013 et 2023 a montré que la dimension la plus souvent évaluée dans les données des DSE est la complétude, suivie de la justesse, de la concordance, de la plausibilité et de l'actualité. Les auteurs y ont ajouté la conformité et le biais, et ils ont constaté qu'il n'existe pas encore de méthode d'évaluation standard [55]. Une autre revue, fondée sur 227 articles, a proposé six dimensions (accessibilité, exactitude, complétude, cohérence, validité contextuelle, actualité) et a désigné la cohérence comme la dimension qui influence toutes les autres [56]. Un exemple montre comment ces dimensions sont liées entre elles. Dans le dossier d'un patient atteint d'insuffisance cardiaque, le poids est noté à l'admission, mais manque à la sortie (problème de complétude). La valeur d'admission a été reprise d'un dossier vieux de deux ans (problème d'actualité). De surcroît, elle est exprimée en livres dans un système et en kilogrammes dans un autre (problème de cohérence). Le calcul automatique de la dose d'un diurétique ou du bilan hydrique part alors d'une donnée erronée, bien que chaque programme fonctionne correctement.

Le chapitre 3 revient sur les méthodes pratiques de contrôle de la qualité.

2.2 La théorie de l'information appliquée à la médecine

2.2.1 Quantité d'information et bit

La théorie mathématique de l'information, formulée par Claude Shannon au milieu du XXe siècle, part d'une idée simple : un message apporte de l'information dans la mesure où il réduit l'incertitude. Un événement certain n'apprend rien de nouveau. Un événement rare surprend beaucoup. La quantité d'information d'un résultat de probabilité p est :

$I = -\log_2 p$ (bits)

Le bit (*binary digit*) est la quantité d'information obtenue lorsqu'on apprend l'issue d'un choix entre deux possibilités également probables. Les fonctions de la théorie de l'information (entropie, entropie relative, information mutuelle) s'appliquent directement aux tests diagnostiques, car l'état de maladie d'un patient et le résultat du test sont des variables aléatoires [57].

Exemple chiffré. Un patient est testé pour une maladie. Si, avant le test, la probabilité de la maladie est de 50 % (0,5), apprendre que le diagnostic est positif apporte $I = -\log_2 0{,}5 = 1$ bit. Si la maladie a une prévalence de 1 % dans la population testée, apprendre que le patient est malade apporte $-\log_2 0{,}01 \approx 6{,}64$ bits, alors qu'apprendre qu'il est sain n'apporte que $-\log_2 0{,}99 \approx 0{,}014$ bit. Pour une prévalence de 0,1 %, un résultat vrai positif représente près de 10 bits (figure 2.2). Intuitivement, apprendre qu'un jeune sans facteur de risque est atteint d'une maladie rare est surprenant, alors que la confirmation qu'il ne l'a pas ne change presque rien.

2.2.2 L'entropie

L'entropie est l'information moyenne attendue d'une variable, c'est-à-dire son incertitude avant l'observation :

$$H = -\Sigma\, p \cdot \log_2 p$$

Pour une variable binaire (malade / sain) de probabilités p et 1 - p, l'entropie est maximale, égale à 1 bit, pour p = 0,5. Pour p = 0,01, $H = -0{,}01 \cdot \log_2 0{,}01 - 0{,}99 \cdot \log_2 0{,}99 \approx 0{,}066 + 0{,}014 \approx 0{,}08$ bit (figure 2.1). La différence a un sens clinique direct. Quand la prévalence est de 50 %, le clinicien ne sait presque rien avant le test. Quand elle est de 1 %, la réponse "sain" est presque certaine d'avance, et l'incertitude restante est faible.

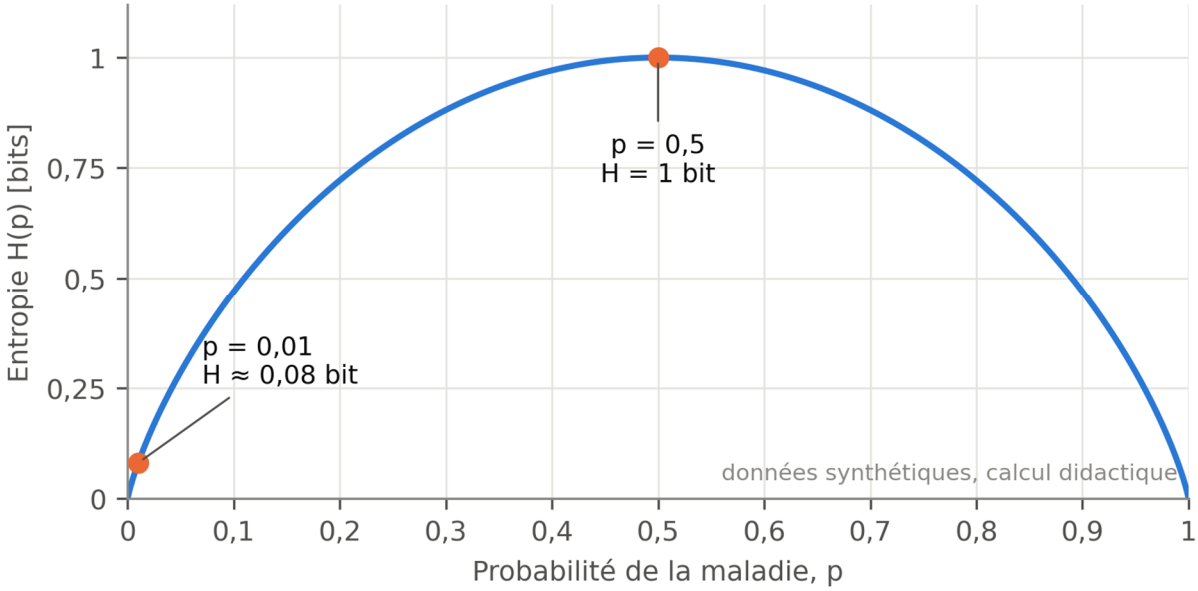


Figure 2.1. Entropie d'une variable binaire $H(p) = -p \cdot \log_2 p - (1-p) \cdot \log_2(1-p)$, avec les points p = 0,5 (1 bit) et p = 0,01 (≈ 0,08 bit). Calcul didactique, données synthétiques.

La redondance est la différence entre la capacité maximale d'un code et l'information qu'il transmet effectivement. En médecine, elle joue un rôle protecteur. Le nom du patient, sa date de naissance et son code numérique personnel répètent en partie la même identité, de sorte qu'une erreur dans un champ peut être décelée à partir des autres. Le bruit est toute perturbation qui altère le message en chemin : un artefact de mouvement sur l'ECG, un chiffre mal transcrit, une phrase mal entendue lors de la relève. Une redondance bien conçue permet de détecter, et parfois de corriger, les erreurs dues au bruit.

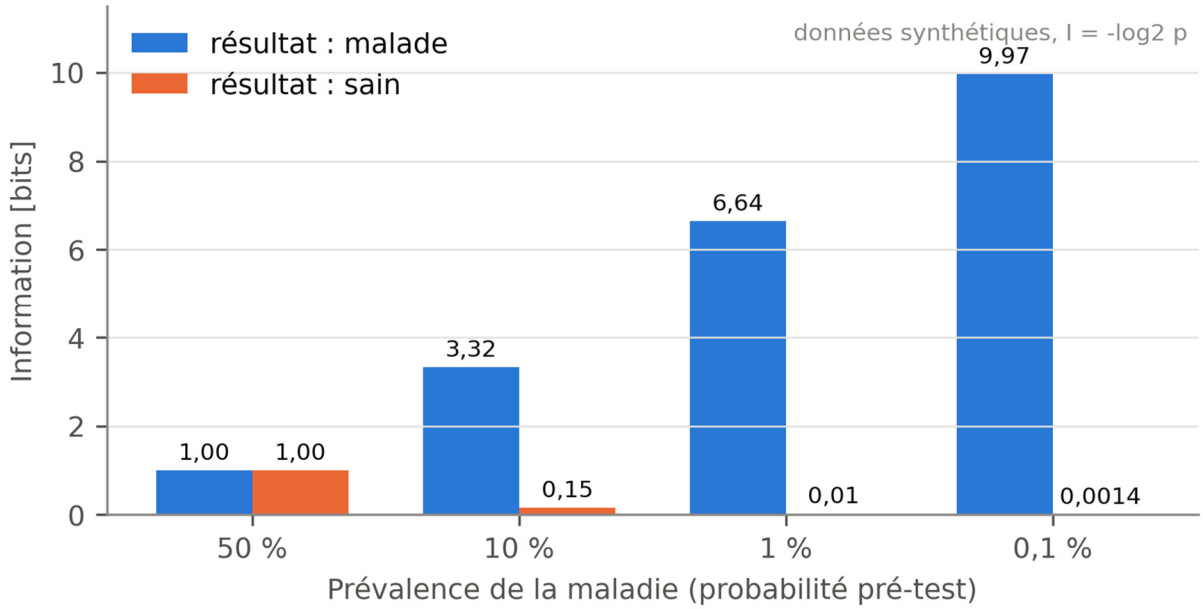


Figure 2.2. Quantité d'information (bits) apportée par la connaissance de l'état réel du patient, pour quatre valeurs de la prévalence de la maladie. Un résultat "malade" apporte d'autant plus de bits que la maladie est rare, alors qu'un résultat "sain" en apporte de moins en moins. Calcul didactique, I = -log2 p, données synthétiques.

2.2.3 L'information apportée par un test diagnostique

Le raisonnement diagnostique classique part de la probabilité pré-test (estimée à partir de l'anamnèse, de l'examen clinique et de la prévalence) et aboutit, grâce au théorème de Bayes, à la probabilité post-test, à l'aide de la sensibilité et de la spécificité du test [58]. Les rapports de vraisemblance (*likelihood ratios*) expriment dans quelle mesure un résultat modifie la probabilité de la maladie et, combinés à la probabilité pré-test, permettent d'estimer la probabilité post-test [59].

La théorie de l'information ajoute une question complémentaire : de combien l'incertitude diminue-t-elle, en moyenne, après le test ? La mesure en est l'information mutuelle, c'est-à-dire l'entropie avant le test moins l'entropie moyenne après le test. La sensibilité et la spécificité, prises séparément, ne mesurent pas explicitement cette réduction, alors que l'information mutuelle l'exprime par un seul nombre [60]. Par exemple, une échographie pour le cancer de l'endomètre, d'une sensibilité de 94 % et d'une spécificité de 40 %, ne réduit l'incertitude diagnostique que de 10 %, un résultat que les deux pourcentages pris isolément masquent [60].

Exemple calculé. Soit un test dont la sensibilité et la spécificité sont de 90 %, appliqué à des patients dont la probabilité pré-test est de 10 % :

- entropie pré-test : $H(0{,}10) \approx 0{,}47$ bit
- 18 % des patients ont un test positif, et chez eux la probabilité post-test est de 50 %, donc H = 1 bit
- 82 % ont un test négatif, avec une probabilité post-test de 1,2 % et $H \approx 0{,}10$ bit
- entropie moyenne après le test : $0{,}18 \cdot 1 + 0{,}82 \cdot 0{,}10 \approx 0{,}26$ bit
- information apportée : 0,47 - 0,26 ≈ 0,21 bit, soit une réduction de l'incertitude d'environ 45 %.

Le même test appliqué à une prévalence de 1 % n'apporte que 0,025 bit, soit une réduction d'environ 31 %. L'observation didactique importante concerne le patient dont le test est positif : la probabilité de la maladie passe de 10 % à 50 %, mais son incertitude individuelle augmente (de

0,47 à 1 bit). Le test a été utile, puisqu'il a fait passer le patient dans une catégorie qui exige des examens complémentaires, mais il n'a pas résolu le diagnostic. Une revue narrative de 20 études ayant utilisé l'entropie dans la décision clinique a constaté qu'elle sert surtout à quantifier l'incertitude et qu'aucune étude ne l'a encore transformée en outil opérationnel de décision validé [61].

2.2.4 Le codage : bit, octet, ASCII, Unicode

L'ordinateur stocke toute information sous forme de suites de bits. Huit bits forment un octet (*byte*), qui peut prendre 2□ = 256 valeurs. La première norme répandue pour le texte, ASCII (*American Standard Code for Information Interchange*), utilisait 7 bits, donc 128 caractères, suffisants pour l'alphabet anglais [62]. Pour l'ensemble des écritures du monde, on a créé Unicode. La version 18.0, publiée le 16 septembre 2026, comprend 172 808 caractères [63]. La forme de codage UTF-8, très répandue dans l'échange de données, représente chaque caractère par 1 à 4 octets et laisse inchangés les caractères ASCII [64].

Pour les noms roumains, ce détail a des conséquences pratiques. Unicode contient deux lettres différentes qui se ressemblent presque : le "s avec virgule souscrite" (ș, code U+0219, introduit explicitement comme ajout pour la langue roumaine) et le "s cédille" (code U+015F), et la même paire existe pour le t (U+021B face à U+0163) [65]. En UTF-8, ș s'écrit sur les octets C8 99, et le s cédille sur C5 9F. La conséquence se voit aussi dans la longueur des chaînes : le nom Mureșan occupe 8 octets en UTF-8, alors que la forme sans diacritiques, Muresan, en occupe 7, car les lettres ASCII sont codées sur un seul octet. Pour un ordinateur, le nom Mureșan écrit avec ș (U+0219) et le même nom écrit avec s cédille (U+015F) sont deux chaînes différentes, bien qu'elles paraissent presque identiques à l'écran. Un même patient peut ainsi apparaître comme deux personnes distinctes si un système (par exemple, un ancien logiciel de laboratoire) utilise la cédille et un autre (le dossier électronique) la virgule.

Exemple

- Un hôpital relie les dossiers d'hospitalisation aux résultats du laboratoire d'après le nom et la date de naissance (sans code numérique personnel, pour des raisons de protection des données).
- Le laboratoire transmet le nom Ștefănescu avec l'initiale codée comme S cédille (U+015E), alors que le dossier électronique l'enregistre avec Ș à virgule (U+0218). La comparaison exacte des chaînes échoue, et les résultats ne sont pas rattachés au dossier.
- La solution : normaliser le texte avant la comparaison (toutes les variantes à cédille sont converties en variante à virgule, et l'on compare éventuellement aussi la forme sans diacritiques) et, lorsque c'est possible, recourir à un identifiant unique commun.

Le problème ne se limite pas à la Roumanie. Pour relier les enregistrements de l'entrepôt de données cliniques de trois hôpitaux universitaires français au registre national des décès, en l'absence d'identifiant commun, un algorithme doté d'un nettoyage avancé des noms et d'une connaissance du système de dénomination a fait passer la sensibilité du couplage de 82,7 % à 93,3 % [66].

2.2.5 La taille des fichiers médicaux

Les ordres de grandeur aident à comprendre les coûts de stockage et de transmission. Un ECG standard de la base publique PTB-XL dure 10 secondes, comporte 12 dérivations, est échantillonné à 500 Hz et stocké avec une précision de 16 bits [67]. On obtient 12 × 5 000

échantillons × 2 octets = 120 000 octets, soit environ 120 ko. Un examen TDM typique comporte des images de 262 144 pixels (matrice 512 × 512), au nombre de 500 environ, ce qui donne à peu près 250 Mo, et une mammographie atteint près de 153 Mo [68]. Les lames d'anatomie pathologique numérisées pèsent habituellement autour de 1 Go chacune [68].

Le génome humain complet (assemblage T2T-CHM13) compte 3,055 milliards de paires de bases [69]. À raison de 2 bits par base (section 2.6), la séquence d'une seule copie tiendrait théoriquement dans environ 764 Mo. En pratique, les données brutes d'un séquençage du génome entier à une couverture d'environ 35× occupent quelque 65 Go en fichiers fastq.gz, et le fichier aligné BAM encore quelque 55 Go [70]. La différence provient de la redondance voulue du séquençage (chaque position est lue des dizaines de fois) et des scores de qualité attachés à chaque base.

Tableau 2.2. Ordres de grandeur de quelques fichiers médicaux. Les valeurs pour l'ECG et pour la séquence de référence sont calculées à partir des paramètres publiés. Source : [67-70].

Type de donnée	Paramètres	Ordre de grandeur
ECG 12 dérivations, 10 s	500 Hz, 16 bits	~120 ko (calculé)
Examen TDM	512 × 512 pixels, ~500 images	~250 Mo
Mammographie	4 images	~153 Mo
Lame numérisée (pathologie)	une image	~1 Go
Génome, une copie de la séquence	3,055 milliards de bases × 2 bits	~764 Mo (calculé)
Génome séquencé, données brutes	couverture ~35×, fastq.gz	~65 Go

2.3 Données, information, connaissance, sagesse

2.3.1 La hiérarchie DIKW

Le modèle DIKW (*data, information, knowledge, wisdom*) ordonne le contenu des systèmes informatiques en quatre niveaux. Les **données** sont des observations brutes. L'**information** désigne des données analysées et organisées de façon à prendre sens, par exemple pour l'affichage ou l'analyse. La **connaissance** résulte de l'interprétation formelle ou informelle de l'information et comprend les résultats des études, mais aussi des heuristiques et des modèles qui peuvent refléter l'expérience ou les préjugés de ceux qui les formulent [11]. La **sagesse** est la capacité d'appliquer cette connaissance à la situation concrète, en pesant les valeurs et le contexte du patient.

L'exemple classique : la pression artérielle de 180/110 mmHg d'un patient est une donnée. Lorsque les données de nombreux patients sont réunies et analysées, on constate que les hypertendus font plus souvent un infarctus du myocarde, et c'est là une connaissance [11]. La figure 2.3 illustre ces niveaux sur un exemple de diabète sucré.

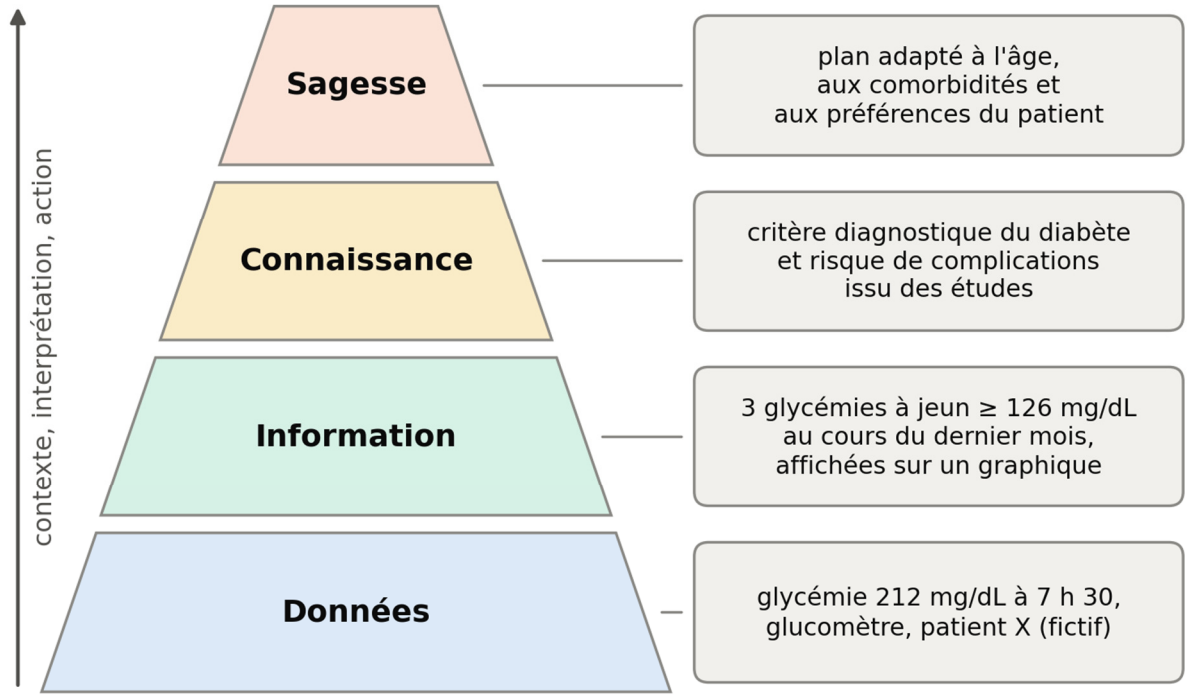


Figure 2.3. Hiérarchie données-information-connaissance-sagesse (DIKW), avec des exemples cliniques tirés du diabète sucré à chaque niveau. Schéma didactique, adapté d'après les définitions de [11].

2.3.2 Les limites du modèle

La pyramide est utile sur le plan didactique, mais elle simplifie la réalité sous au moins trois aspects :

- **les frontières dépendent du niveau d'abstraction**. L'affirmation "le patient est hypertendu" est une interprétation des mesures, mais elle devient une simple donnée pour un système qui calcule le risque cardiovasculaire [11]
- **le flux n'est pas unidirectionnel**. La connaissance existante décide des données recueillies : le médecin qui soupçonne une maladie demande certains examens et pas d'autres
- **l'action est absente**. En surveillance de santé publique, on a proposé d'insérer un niveau d'"intelligence" (connaissance mise en action) entre la connaissance et la sagesse, pour que la communication des données soit suivie d'une réponse [71].

Les graphes de connaissances (*knowledge graphs*) sont un exemple moderne de passage des données à la connaissance, par la mise en relation d'entités biomédicales au sein d'un réseau. L'intégration de données réelles issues des DSE y est toutefois restée limitée [72].

2.3.3 Connaissance explicite et connaissance tacite

La connaissance **explicite** peut être écrite, transmise et vérifiée : un guide de pratique clinique, une posologie, un algorithme de réanimation. La connaissance **tacite** relève de l'expérience et se transmet difficilement par des mots. L'impression clinique au lit du malade ou la "main" du chirurgien en sont des exemples. Une revue systématique accompagnée d'une analyse herméneutique a montré que, dans des situations critiques comme la pandémie de COVID-19, les guides et la connaissance scientifique n'ont pas suffi à orienter l'action, et que la connaissance tacite a servi de mode d'enseignement et d'apprentissage de savoir-faire liés aux valeurs et à

l'éthique [73]. Pour l'informatique, la distinction a une conséquence pratique : seule la connaissance explicite peut être codée directement dans un système.

2.3.4 La représentation des connaissances

Pour être utilisée par l'ordinateur, la connaissance explicite doit être écrite dans un formalisme :

- **les règles de production** ont la forme SI ... ALORS ... Le système MYCIN, prototype des systèmes à base de règles, représentait ainsi les connaissances sur les infections et les antibiotiques [14]. Le standard Arden Syntax organise des règles de type événement-condition-action en modules logiques médicaux (par exemple, une alerte pour l'examen annuel du pied chez le diabétique) [14]
- **les arbres de décision** représentent la succession des choix et des événements incertains, chaque branche étant associée à une probabilité et à une conséquence [58]
- **les ontologies** organisent les types d'entités d'un domaine en une hiérarchie et rendent explicites les relations qui les unissent, sous la forme d'un graphe [14]. La création d'ontologies et de graphes de connaissances reste une direction de recherche majeure [74].

Le choix du formalisme n'est pas neutre. Une analyse de plusieurs technologies de règles (Arden Syntax, Drools, SWRL et d'autres), appliquée aux guides d'antibiothérapie, a montré qu'elles diffèrent par leur expressivité, leur interopérabilité et leur maturité industrielle [75]. Les terminologies et ontologies médicales sont traitées au chapitre 4, et les systèmes d'aide à la décision, au chapitre 13.

2.4 La communication de l'information dans l'acte médical

2.4.1 Le modèle émetteur-canal-récepteur

Shannon a décrit la communication comme une chaîne : une source produit un message, un émetteur le code, le message traverse un canal où intervient le bruit, et le récepteur le décode. Dans la transmission d'un patient lors de la relève (*handover* ou *handoff*), l'émetteur est le médecin qui s'en va, le récepteur est le médecin qui prend le relais, et le canal est l'échange oral, la feuille écrite ou le dossier électronique (figure 2.4). Le bruit prend des formes concrètes. Une méta-synthèse des études qualitatives sur la transmission en soins intensifs a relevé, entre autres, les interruptions et les distractions, l'absence d'informations pertinentes et la communication unidirectionnelle [76]. Une revue systématique de 60 études sur la transmission du patient entre l'équipe d'urgence préhospitalière et le personnel du service des urgences a décrit des facteurs environnementaux, la dégradation et la perte de l'information, le manque de standardisation et une mémorisation défaillante [77].

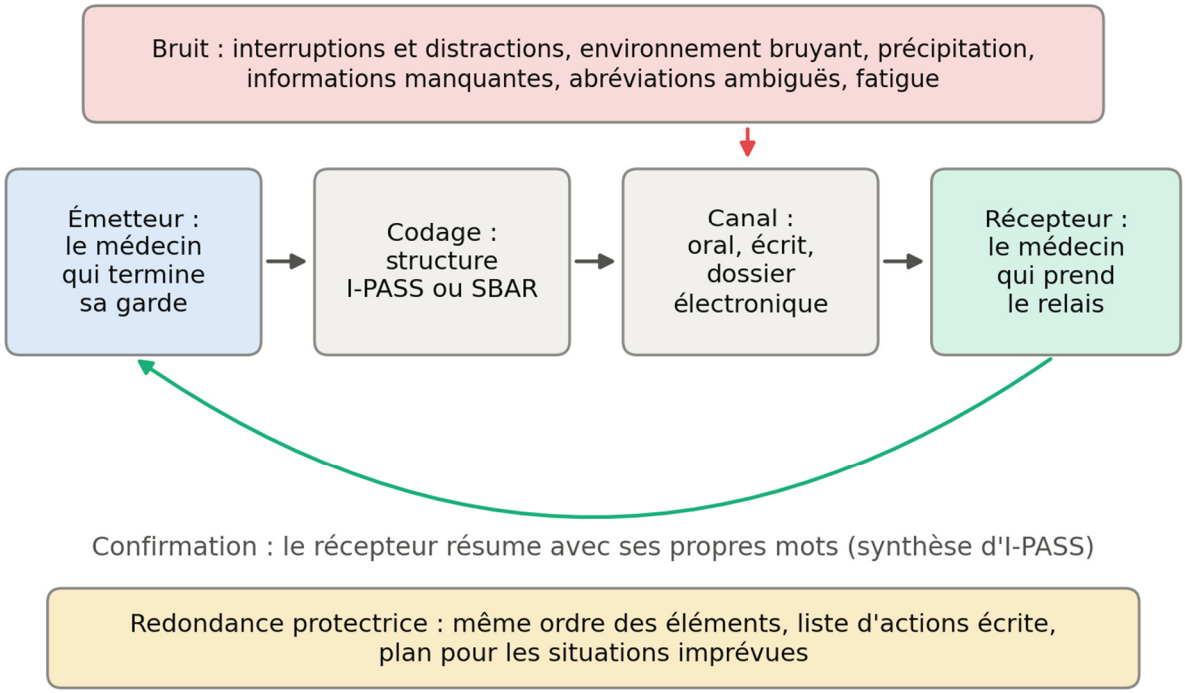


Figure 2.4. Le modèle émetteur-canal-récepteur appliqué à la transmission du patient lors de la relève, avec les sources de bruit et les mécanismes de protection (structure standard, confirmation par un résumé du récepteur). Schéma didactique d'après le modèle de Shannon, avec des sources de bruit tirées de [76,77].

2.4.2 Les erreurs de communication et les protocoles structurés

Les conséquences sont mesurables. Dans un échantillon de dossiers de plaintes pour faute médicale aux États-Unis, des défaillances de communication ont été identifiées dans 49 % des cas, et 40 % de ces défaillances concernaient une transmission ratée. Les informations le plus souvent mal transmises étaient les plans pour les situations imprévues, le diagnostic et la gravité de la maladie [78].

Les protocoles structurés appliquent en fait les principes de la théorie de l'information : un ordre fixe des éléments (code commun à l'émetteur et au récepteur) et une confirmation par le récepteur (redondance). **SBAR** (*Situation, Background, Assessment, Recommendation*) demande de présenter la situation, le contexte, l'évaluation et la recommandation. **I-PASS** (*Illness severity, Patient summary, Action list, Situation awareness and contingency planning, Synthesis by receiver*) ajoute explicitement le plan pour les situations imprévues et la synthèse faite par le récepteur.

Les preuves réunies en 2020-2025 sont encourageantes, mais inégales. Dans une étude prospective menée dans 32 hôpitaux auprès de 2 735 internes, la mise en œuvre d'I-PASS a été associée à une baisse de 47 % des événements indésirables signalés liés à la transmission (les événements majeurs sont passés de 1,7 à 0,9 par personne-année). La proportion de transmissions orales comportant les cinq éléments clés est passée de 20 % à 66 % [79]. Une revue systématique de 2025 a conclu qu'I-PASS réduit les erreurs médicales et les événements indésirables, avec un niveau de certitude des preuves modéré, alors que, pour SBAR, les preuves d'une amélioration des résultats de sécurité sont de faible certitude [80]. Pour SBAR, une autre revue a montré que l'application fidèle du protocole et une plus grande clarté de la communication ont été observées surtout dans les études menées en salle de cours. Des effets sur les patients n'ont été constatés que dans des interventions complexes, à plusieurs composantes [81].

Le canal peut aussi être électronique. Une revue systématique de 9 publications sur la transmission au moyen de plateformes électroniques a suggéré que celle-ci peut améliorer la complétude de l'information transmise. On ignore toutefois si ce gain se traduit par de meilleurs résultats pour les patients [82]. Un formulaire électronique de transmission reste donc un soutien pour la structure, et non un substitut à l'échange direct et à la confirmation faite par le récepteur.

Exemple

Transmission I-PASS, patient fictif, service de médecine interne, 20 h.

- I (gravité) : "Patient instable, sous surveillance rapprochée."
- P (résumé) : "Homme de 68 ans, pneumonie communautaire, deuxième jour d'antibiotique."
- A (actions) : "Réévaluation de la saturation à 22 h, hémocultures à vérifier demain matin."
- S (situations imprévues) : "Si la SpO2 descend en dessous de 90 % sous oxygène au masque, on appelle le médecin de garde d'anesthésie-réanimation (ATI)."
- S (synthèse du récepteur) : le médecin qui prend le relais reformule avec ses propres mots, et les divergences sont clarifiées sur-le-champ.

2.5 La médecine fondée sur les preuves au format numérique

2.5.1 Niveaux de preuve et certitude des preuves

La médecine fondée sur les preuves (*evidence-based medicine*, EBM) associe les meilleures preuves disponibles à l'expérience du clinicien et aux préférences du patient. Traditionnellement, les preuves sont hiérarchisées selon le plan d'étude : revues systématiques et méta-analyses, essais contrôlés randomisés, études de cohorte, études cas-témoins, séries de cas, opinion d'experts. L'approche GRADE (*Grading of Recommendations Assessment, Development and Evaluation*) évalue cependant la certitude des preuves pour chaque critère de jugement séparément, à partir d'une question formulée selon la structure PICO (population, intervention, comparateur, critère de jugement), et elle distingue le jugement porté sur les preuves du passage à la recommandation [83].

2.5.2 Synthèses de preuves et revues systématiques vivantes

Le clinicien ne peut pas lire toutes les études primaires. On recourt donc aux synthèses : revues systématiques, revues de portée (*scoping reviews*), cartes des preuves, revues de revues. Les revues rapides et les revues vivantes sont des modes de travail applicables à plusieurs types de synthèse [84]. Une revue systématique vivante (*living systematic review*) est mise à jour en continu, à mesure que paraissent de nouvelles études. Elle est surtout utile lorsque les preuves apparaissent vite, qu'elles sont incertaines et que de nouveaux résultats peuvent modifier la pratique [85]. Le même principe s'applique aux guides vivants (*living guidelines*). Leur méthodologie comprend la décision d'adopter ou non l'approche vivante pour un guide donné, la veille de la littérature et la fréquence des recherches, ainsi que la sortie des recommandations du mode vivant lorsque le sujet se stabilise [86].

Au point de soins, le médecin consulte habituellement des sources de synthèse déjà évaluées : la base Cochrane de revues systématiques, BMJ Best Practice, UpToDate et des produits similaires [13].

Tableau 2.3. Types de sources de preuves et leur utilisation. Source : [13,84-86].

Type de source	Ce qu'elle apporte	Quand l'utiliser
Étude primaire (ECR, cohorte)	Le résultat d'une seule recherche	Questions nouvelles, sans synthèse
Revue systématique	Synthèse de toutes les études sur une question	Évaluer l'effet d'une intervention
Revue systématique vivante	Synthèse mise à jour en continu	Domaines où les preuves évoluent vite
Guide clinique (GRADE)	Recommandations assorties de la certitude des preuves	Décisions de pratique
Ressource de synthèse au lit du malade	Résumés cliniques actualisés	Questions ponctuelles pendant la consultation

2.5.3 Guides cliniques computables

Un guide publié sous forme de texte (PDF) doit être lu et interprété par chaque médecin. Un guide **computable** est traduit sous une forme que le système informatique peut exécuter sur les données du patient. On a proposé, par exemple, un format fondé sur des ressources FHIR qui représente la recommandation, sa force, la certitude des preuves et le lien avec les revues systématiques sur lesquelles elle s'appuie [87]. Les Centres pour le contrôle et la prévention des maladies (CDC) des États-Unis ont traduit les recommandations de dépistage des cancers du col de l'utérus, du sein et colorectal en guides computables et en outils d'aide à la décision. L'outil consacré au cancer du col a été intégré au DSE et testé en clinique pendant 6 mois. Les deux autres n'ont été testés qu'avec des données synthétiques [88]. Les auteurs notent que la généralisation dépend de la saisie structurée des données, autrement dit précisément de la qualité des données discutée à la section 2.1.

2.5.4 L'automatisation de la synthèse des preuves

Une revue systématique est coûteuse et lente. Un rapport de l'AHRQ de décembre 2024, conçu comme une revue rapide vivante assortie d'une carte des preuves, a inclus 56 études ayant évalué des outils d'apprentissage automatique pour différentes étapes [89]. Les résultats varient fortement d'une étape à l'autre :

- les outils d'identification des essais randomisés ont atteint une sensibilité (*recall*) de 96-100 % (médiane 98,5 %)
- les outils de tri des résumés ont eu une sensibilité médiane de 93 % et ont réduit le volume de travail d'une médiane de 50 %
- l'extraction automatique des données a été correcte dans des proportions allant de 0 à 99 % (médiane 10 %)
- l'évaluation du risque de biais a montré un accord modeste avec les évaluateurs humains (kappa pondéré médian 0,16) [89].

Conclusion du rapport : l'identification des essais randomisés et la priorisation des résumés se rapprochent d'un usage courant sous supervision humaine, alors que la recherche et l'extraction des données ne sont pas encore assez fiables [89]. Une étude préenregistrée menée avec GPT-4 a observé des performances modérées pour l'extraction des données et variables pour le tri, et les auteurs recommandent une grande prudence dans l'utilisation des grands modèles de langage pour les revues systématiques [90].

2.6 L'information en biologie

2.6.1 L'ADN comme code

La molécule d'ADN utilise un alphabet de quatre bases (A, C, G, T). Selon la formule de Shannon appliquée à quatre symboles équiprobables, chaque base peut porter au plus log2 4 = 2 bits. Le génome humain complet, de 3,055 milliards de paires de bases [69], contient donc, en ordre de grandeur, environ 6 milliards de bits pour une seule copie. Ce chiffre est une limite supérieure théorique. La séquence réelle n'est pas aléatoire (elle contient des régions répétitives), si bien que l'information effective est moindre, et les programmes de compression exploitent précisément cette redondance.

La densité de l'ADN comme support de stockage est très élevée. On l'estime à environ 10^{1}□ bits par centimètre cube, ce qui explique les recherches sur l'utilisation de l'ADN synthétique pour l'archivage de données numériques [91]. La même logique relie l'informatique à la biologie. La génomique produit des volumes de données comparables à ceux de l'imagerie, et leur analyse exige les méthodes de stockage et de gestion présentées au chapitre 3.

2.6.2 Note sur l'information quantique

Les ordinateurs classiques travaillent avec des bits qui valent soit 0, soit 1. Le calcul quantique utilise des **qubits** qui, par superposition, peuvent se trouver dans une combinaison des états 0 et 1 et qui, par intrication quantique (*entanglement*), peuvent être corrélés entre eux. Une revue de 2024 décrit des applications explorées dans la découverte de médicaments, la modélisation moléculaire, la génomique et le diagnostic. La même revue énumère les limites actuelles : le matériel, le passage à l'échelle et l'intégration dans l'environnement clinique [92]. Pour l'étudiant de première année, la conclusion prudente est qu'au 27 septembre 2026, le calcul quantique reste un domaine de recherche, et non un outil clinique de routine.

Pièges fréquents

- Confondre donnée et information : une valeur sans unité, sans date et sans méthode de mesure ne peut pas être interprétée correctement.
- Traiter une variable ordinale codée par des chiffres comme une variable numérique (la moyenne du "degré de douleur").
- Croire qu'un résultat positif "clarifie" toujours le diagnostic. À faible prévalence, un test positif peut laisser le patient dans une incertitude plus grande qu'avant.
- Négliger les diacritiques : ş à virgule (U+0219) et s cédille (U+015F) sont des caractères différents pour l'ordinateur, et le couplage des enregistrements d'après le nom peut échouer.
- Confondre bit et octet : 1 octet = 8 bits, et les débits de transmission s'expriment habituellement en bits par seconde.
- Transmettre un patient sans plan pour les situations imprévues et sans confirmation du récepteur.

À retenir

- Une donnée clinique comporte cinq composantes : patient, paramètre, valeur, moment, méthode.
- La quantité d'information d'un résultat est I = -log2 p. Les événements rares apportent davantage de bits.
- L'entropie d'une variable binaire est maximale (1 bit) pour p = 0,5 et tombe à environ 0,08 bit pour p = 0,01.
- L'information apportée par un test est l'entropie pré-test moins l'entropie moyenne post-test.
- Un ECG de 10 s occupe environ 120 ko, une TDM environ 250 Mo, les données brutes d'un génome environ 65 Go.
- DIKW est un modèle didactique utile, dont les frontières dépendent du contexte.
- I-PASS s'appuie sur des preuves de certitude modérée pour la réduction des erreurs de transmission.
- L'automatisation de la synthèse des preuves fonctionne bien pour le tri, mais pas encore pour l'extraction des données.

Questions d'autoévaluation

1. (QCS) Une maladie a une prévalence de 1 % dans la population testée. La quantité d'information apportée par le fait d'apprendre qu'un patient est atteint de la maladie est d'environ :

a) 0,01 bit
b) 0,08 bit
c) 1 bit
d) 6,6 bits
e) 100 bits

2. (QCS) L'entropie d'une variable binaire (malade / sain) est maximale lorsque la probabilité de la maladie est égale à :

a) 0
b) 0,01
c) 0,25
d) 0,5
e) 1

3. (QCS) Laquelle des variables suivantes est ordinale ?

a) le groupe sanguin
b) la classe fonctionnelle NYHA
c) la créatinine sérique
d) le service d'hospitalisation
e) le poids corporel

4. (QCM) À propos du codage des caractères dans les données médicales, quelles affirmations sont exactes ?

a) ASCII utilisait 7 bits et pouvait représenter 128 caractères
b) UTF-8 utilise de 1 à 4 octets pour un caractère
c) ş à virgule (U+0219) et s cédille (U+015F) sont le même caractère Unicode
d) la différence entre ş et s cédille peut empêcher le couplage des enregistrements d'un patient d'après le nom
e) UTF-8 modifie les codes des caractères ASCII

5. (QCM) À propos de la transmission du patient lors de la relève, quelles affirmations sont exactes ?

a) dans I-PASS, le dernier élément est la synthèse faite par le récepteur
b) les interruptions et les distractions sont des sources de bruit décrites dans la transmission en soins intensifs
c) une revue systématique de 2025 a trouvé une certitude modérée des preuves en faveur d'I-PASS pour la réduction des erreurs médicales
d) SBAR a démontré, avec une certitude élevée, une réduction de la mortalité
e) le plan pour les situations imprévues ne fait pas partie d'I-PASS

CHAPITRE 3

Chapitre 3. Données structurées, fichiers et bases de données

Objectifs

À la fin de ce chapitre, l'étudiant pourra :

- comparer les formats de fichier CSV, JSON et XML sur un même enregistrement de patient fictif
- construire un diagramme entité-association simple et y repérer les clés primaires et étrangères
- normaliser jusqu'à la troisième forme normale une table clinique redondante
- écrire des requêtes SQL élémentaires avec SELECT, WHERE, JOIN et GROUP BY
- expliquer le rôle du modèle commun de données OMOP et des principes FAIR dans la recherche sur les données de santé
- choisir un outil de collecte électronique des données adapté à une étude, à partir de critères de validation, d'audit et de conformité

3.1 Fichiers et formats

3.1.1 Fichier, format, métadonnées

Un fichier est une suite d'octets portant un nom, conservée sur un support de stockage. Le format est la convention qui indique comment ces octets doivent être lus. Les mêmes octets peuvent représenter un texte, une image ou un signal, et un programme qui ne connaît pas le format ne peut pas les interpréter. Les métadonnées sont des "données sur les données" : qui a créé le fichier, quand, avec quel appareil, dans quelles unités sont exprimées les valeurs, ce que signifie chaque colonne. Sans métadonnées, une table comportant la colonne `tas` et la valeur 145 ne dit pas s'il s'agit de la pression systolique en mmHg, ni si elle a été mesurée au cabinet ou à domicile.

3.1.2 Formats texte : CSV, JSON, XML

Les formats texte peuvent être ouverts avec n'importe quel éditeur et sont, en général, les plus portables. La portabilité signifie que les données peuvent être lues par d'autres programmes, et même après de nombreuses années, sans le programme qui les a créées.

CSV (*comma-separated values*) est une table écrite sous forme de texte : chaque ligne est un enregistrement, et les valeurs sont séparées par une virgule (ou par un point-virgule, dans les configurations régionales qui utilisent la virgule comme séparateur décimal). La première ligne contient habituellement les noms des colonnes. Le même enregistrement d'un patient fictif se présente ainsi :

- `id_pacient,data_nasterii,sex,tas,tad,fumator`
- `P0001,1980-03-14,F,128,82,non`

JSON (*JavaScript Object Notation*) organise les données en paires nom-valeur, qui peuvent être imbriquées. Le même enregistrement, enrichi de deux mesures, devient :

```
{"id_pacient": "P0001", "data_nasterii": "1980-03-14", "sex": "F",
 "tensiune": [{"data": "2026-09-20", "tas": 128, "tad": 82},
              {"data": "2026-09-27", "tas": 134, "tad": 85}]}
```

La différence saute aux yeux. Le CSV a une structure plate, avec une seule valeur par cellule, si bien que deux mesures de la pression artérielle exigent soit deux lignes pour le même patient, soit des colonnes supplémentaires (tas_1, tas_2). Le JSON accepte des listes et des objets à l'intérieur d'autres objets, il est donc semi-structuré. **XML** (*eXtensible Markup Language*) exprime la même hiérarchie par des balises d'ouverture et de fermeture (`<tas>128</tas>`). Il est plus "verbeux", mais permet une validation stricte par rapport à un schéma. Le standard HL7 FHIR Release 5 permet de représenter les ressources cliniques en XML, en JSON et en Turtle [93]. L'interopérabilité est traitée en détail au chapitre 8.

Deux détails de format provoquent souvent des erreurs lors de l'échange de fichiers CSV. Le premier est le codage des caractères. Un fichier enregistré en UTF-8 et ouvert par un programme qui suppose un autre codage affiche mal les diacritiques, et les noms des patients ne peuvent plus être comparés (chapitre 2). Le second est le séparateur décimal. Dans la configuration roumaine, la virgule sépare les décimales (37,5 °C), de sorte qu'un fichier aux valeurs séparées par des virgules peut être lu avec des colonnes décalées. La solution consiste à convenir explicitement, avant l'échange, du codage (UTF-8), du séparateur de champs et du format des dates (AAAA-MM-JJ, qui se trie correctement même comme texte).

3.1.3 Formats binaires et portabilité

Les formats binaires stockent les valeurs directement dans la représentation interne de l'ordinateur. Ils sont compacts et se lisent vite, mais ils sont illisibles sans le programme adéquat. Les fichiers DICOM d'imagerie (chapitre 7), les fichiers de signal ECG ou EEG (chapitre 6) et les formats génomiques compressés en sont des exemples médicaux. Un format binaire ouvert et documenté est beaucoup plus sûr à long terme qu'un format propriétaire, que seul le fabricant peut lire. Pour l'archivage des données d'une étude, la règle pratique est la suivante : les données dans un format ouvert (CSV codé en UTF-8, chapitre 2), accompagnées d'un dictionnaire de données (section 3.6) et d'une description des traitements effectués.

3.2 La modélisation des données

3.2.1 Le modèle entité-association

Avant de construire une base de données, on décrit les "choses" (entités) qui nous intéressent et la manière dont elles sont liées. Dans un service de consultations externes fictif, les entités sont : **Patient**, **Consultation**, **Diagnostic**, **Prescription** et **Médicament**. Chaque entité possède des attributs (le patient a une date de naissance et un sexe, la consultation a une date). Les associations ont une cardinalité :

- un patient peut avoir de nombreuses consultations, mais chaque consultation appartient à un seul patient (association un-à-plusieurs, 1:N)
- lors d'une consultation, on peut poser plusieurs diagnostics et émettre plusieurs prescriptions (1:N)

- un médicament peut figurer dans de nombreuses prescriptions (1:N de Médicament vers Prescription).

La figure 3.1 présente le diagramme entité-association de cette base de données. Les données cliniques regorgent d'associations un-à-plusieurs, qu'un tableur gère mal. Une base de données relationnelle les gère naturellement [94].

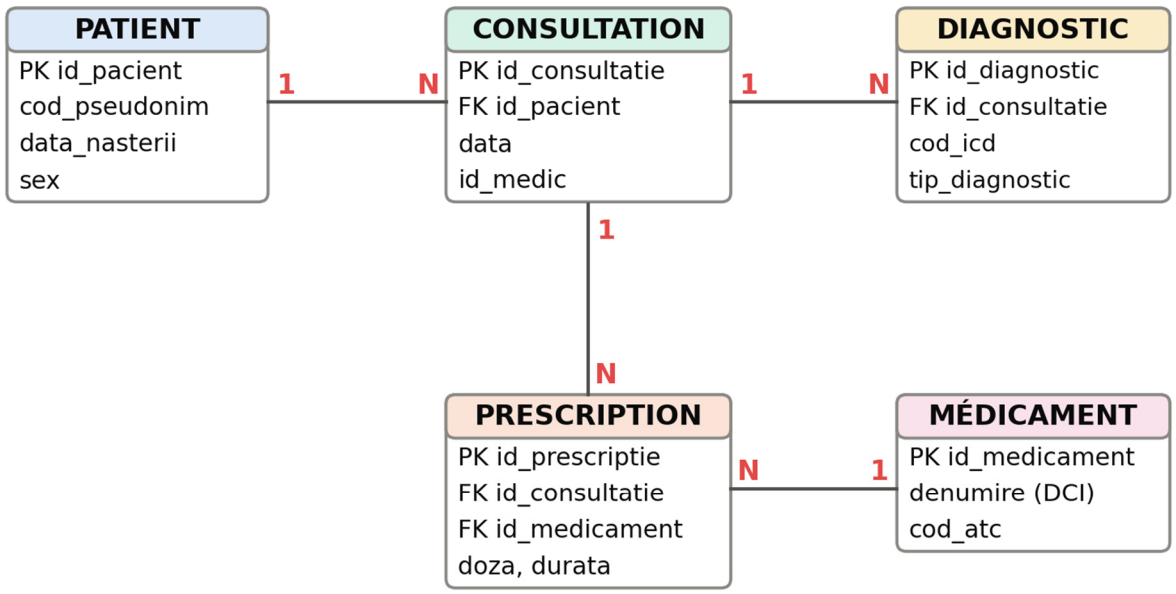


Figure 3.1. Diagramme entité-association de la base de données fictive d'un service de consultations externes : cinq entités, les clés primaires (PK), les clés étrangères (FK) et les cardinalités 1:N. Schéma didactique.

3.2.2 Clés primaires et clés étrangères

Chaque table possède une **clé primaire** (*primary key*), c'est-à-dire un attribut ou un groupe d'attributs qui identifie de manière unique chaque ligne. Sa valeur ne peut pas manquer, et une table n'a qu'une seule clé primaire [95]. Le nom du patient n'est pas une bonne clé, car deux patients peuvent porter le même nom, et un nom peut changer ou être écrit différemment (voir le problème des diacritiques au chapitre 2). On utilise donc un identifiant artificiel, par exemple `id_pacient`.

Une **clé étrangère** (*foreign key*) est une colonne qui doit désigner une ligne existante d'une autre table [95]. La colonne `id_pacient` de la table Consultation est une clé étrangère vers la table Patient. Le système refuse ainsi l'enregistrement d'une consultation pour un patient inexistant. Le fait que chaque enregistrement possède une clé empêche aussi d'ajouter deux fois le même patient, et les champs ont un type de données déclaré, ce qui bloque la saisie de valeurs dénuées de sens [94].

3.2.3 La normalisation

La normalisation est le processus par lequel une table redondante est décomposée en tables plus petites, liées par des clés, de sorte que chaque fait ne soit écrit qu'une seule fois. On part souvent d'une feuille de calcul du type suivant :

Tableau 3.1. Table non normalisée contenant des données fictives de consultations externes. Chaque ligne répète les données du patient, et une cellule contient plusieurs diagnostics. Données fictives.

id_consultatie	patient (nom, année de naissance)	date	diagnostics	médicament prescrit
C01	Patient A, 1958	2026-09-02	I10, E11.9	metformine 850 mg
C02	Patient A, 1958	2026-09-23	I10	amlodipine 5 mg
C03	Patient B, 1990	2026-09-23	J06.9	paracétamol 500 mg

Cette table présente trois défauts. Les données du patient A sont écrites deux fois, et une correction faite dans une seule ligne crée une contradiction (anomalie de mise à jour). La colonne "diagnostics" contient une liste, et non une valeur. Enfin, la dose et le nom du médicament sont mêlés dans le même texte.

- **Première forme normale (1NF)** : chaque cellule contient une seule valeur atomique, et les groupes répétitifs sont déplacés dans des lignes distinctes. Les diagnostics de la consultation C01 deviennent deux lignes dans une table Diagnostic.
- **Deuxième forme normale (2NF)** : en outre, chaque attribut qui ne fait pas partie de la clé dépend de la clé primaire entière, et pas seulement d'une partie de celle-ci. Si une table a pour clé composée (id_consultatie, cod_ICD), le libellé du diagnostic ne dépend que de cod_ICD et doit être déplacé dans une table de nomenclature.
- **Troisième forme normale (3NF)** : en outre, aucun attribut qui ne fait pas partie de la clé ne dépend d'un autre attribut qui n'en fait pas partie. Dans la table Consultation, l'année de naissance dépend du patient, et non de la consultation, elle est donc déplacée dans la table Patient.

On obtient la structure de la figure 3.1 : chaque patient apparaît une fois, chaque médicament une fois, et les consultations, les diagnostics et les prescriptions y renvoient par des clés. Pour accélérer les analyses, les entrepôts de données renoncent parfois délibérément à une partie de la normalisation (section 3.4), mais la base de données dans laquelle les données sont saisies doit être normalisée.

La normalisation a aussi un coût. Pour reconstituer la "fiche" d'une consultation, le système doit relier de nouveau quatre ou cinq tables, et les requêtes s'allongent. C'est pourquoi, en pratique, on conserve une base normalisée pour la saisie et la mise à jour des données et l'on construit séparément des vues (*views*) ou des tables dérivées pour les rapports. La règle didactique reste simple : un fait se modifie en un seul endroit.

3.3 Systèmes de gestion de bases de données relationnelles et SQL

3.3.1 SGBD et langage SQL

Un système de gestion de base de données (SGBD, en anglais *database management system*) est le programme qui stocke les tables, vérifie les règles d'intégrité, gère l'accès simultané de plusieurs utilisateurs et répond aux requêtes. PostgreSQL, MySQL, Microsoft SQL Server et Oracle sont des exemples de SGBD relationnels. Tous utilisent le langage SQL (*Structured Query Language*), qui permet de sélectionner avec précision des sous-ensembles de patients parmi un très grand nombre de champs [94]. Sur la base fictive de la figure 3.1, quelques requêtes se présentent ainsi :

- sélection avec filtre : `SELECT id_pacient, data_nasterii FROM pacient WHERE sex = 'F'`

- jointure entre tables : `SELECT p.id_pacient, c.data FROM pacient p JOIN consultatie c ON c.id_pacient = p.id_pacient WHERE c.data >= '2026-09-01'`
- regroupement et comptage : `SELECT d.cod_icd, COUNT(*) AS nr FROM diagnostic d GROUP BY d.cod_icd ORDER BY nr DESC`

La première requête renvoie les patientes. La deuxième relie chaque consultation à son patient (JOIN) et ne garde que les consultations de septembre 2026. La troisième compte combien de fois chaque diagnostic a été posé, autrement dit elle produit une table de fréquences. La même logique sous-tend les tableaux croisés dynamiques des tableurs (TP 3), mais SQL l'applique à des tables de plusieurs millions de lignes.

3.3.2 Transactions ACID et intégrité

Une transaction est un groupe d'opérations qui doivent être traitées comme un tout. L'exemple clinique type est la délivrance d'un médicament : la sortie du stock de la pharmacie et l'enregistrement de l'administration au patient doivent se produire ensemble ou pas du tout. Les propriétés ACID décrivent les garanties d'un SGBD [95] :

- **atomicité** : toutes les opérations de la transaction s'achèvent comme une unité, ou aucune, et après une panne du système il ne subsiste aucun résultat partiel
- **cohérence** : les données respectent en permanence les contraintes d'intégrité, et la transaction qui les enfreindrait est annulée
- **isolation** : les effets d'une transaction ne sont pas visibles pour les autres transactions simultanées avant sa validation
- **durabilité** : une fois validée, la modification subsiste même après une panne du système.

Les contraintes d'intégrité sont des règles déclarées dans la base de données : champ obligatoire (NOT NULL), valeurs uniques, intervalle autorisé (par exemple, un CHECK qui exige que la pression systolique soit comprise entre 50 et 300 mmHg) et clés étrangères. Elles constituent la première ligne de défense de la qualité des données (section 3.6).

3.4 Bases NoSQL, entrepôts et lacs de données, mégadonnées

3.4.1 Les modèles NoSQL

Le terme NoSQL (*not only SQL*) recouvre les bases de données qui n'organisent pas les données en tables relationnelles. Les principaux modèles sont au nombre de quatre :

- **document** : chaque enregistrement est un document JSON, à structure flexible (par exemple, un résumé de patient avec tous les diagnostics imbriqués)
- **clé-valeur** : une clé unique désigne une valeur opaque (utile pour les sessions d'utilisateur ou les caches d'images)
- **colonnes** (familles de colonnes) : les données sont stockées par colonnes, ce qui est efficace pour de très grands volumes et les requêtes analytiques
- **graphe** : des nœuds (patient, diagnostic, médicament) et des arêtes (a le diagnostic, reçoit le médicament), efficace pour les relations multiples et les chemins dans un réseau.

Une revue comparative a montré que les bases NoSQL de type document, colonne et graphe gèrent efficacement les caractéristiques propres aux données standardisées des dossiers de santé électroniques (DSE, *electronic health record*, EHR), surtout dans les systèmes distribués [96]. Les modèles de type graphe sont de plus en plus utilisés pour les données cliniques et biomédicales, par exemple pour le phénotypage des maladies et la prédiction du risque, bien que l'intégration

des données issues des DSE y soit restée limitée [72]. La figure 3.2 compare les trois principaux modèles sur un même patient fictif.

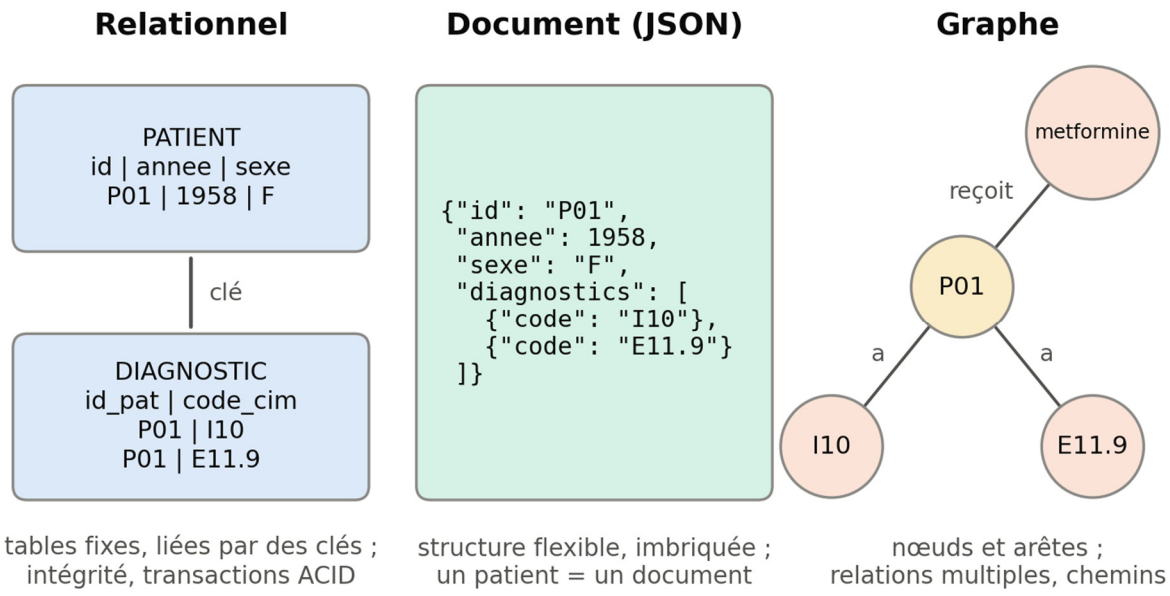


Figure 3.2. Un même patient fictif représenté dans trois modèles de données : relationnel (tables liées par des clés), document (JSON imbriqué) et graphe (nœuds et relations). Schéma didactique.

Le choix n'est pas "SQL ou NoSQL" : il dépend de la tâche. Les transactions cliniques (prescriptions, hospitalisations, facturation) ont besoin des garanties ACID et de l'intégrité référentielle du modèle relationnel. Les recherches en texte libre et les requêtes sur de très grands volumes peuvent tirer parti d'autres modèles. Au centre hospitalier universitaire de Bordeaux, un entrepôt i2b2 de 2,5 millions de patients et de plus de 3 milliards d'observations a été transféré de la base relationnelle vers le moteur NoSQL Elasticsearch. Les temps de requête ont diminué, avec des gains importants pour les recherches en texte libre, et l'espace disque occupé a été moindre. En revanche, le modèle a dû être adapté, car Elasticsearch ne peut pas effectuer de jointures entre des index différents [97].

Tableau 3.2. Comparaison des modèles de bases de données pour les données de santé. Source : synthèse d'après [72,94,96,97].

Modèle	Unité de stockage	Point fort	Exemple d'utilisation
Relationnel	Table avec lignes et colonnes	Intégrité, transactions ACID	Prescriptions, hospitalisations
Document	Document JSON	Structure flexible	Résumé de patient, messages FHIR
Clé-valeur	Paire clé-valeur	Accès direct par la clé	Sessions, cache
Colonnes	Famille de colonnes	Grands volumes, analyses par colonnes	Séries de signaux, journaux
Graphe	Nœuds et arêtes	Relations multiples	Graphes de connaissances

3.4.2 Entrepôts de données et lacs de données

Les systèmes de routine (dossier électronique, laboratoire, pharmacie) sont optimisés pour les transactions : de nombreuses petites opérations, portant sur un seul patient. Les analyses de

recherche demandent autre chose : des requêtes lourdes, sur de nombreux patients et sur de longues périodes. C'est pourquoi les données sont copiées périodiquement dans un **entrepôt de données** (*data warehouse*), où elles sont nettoyées, intégrées et organisées pour l'analyse. Un entrepôt de données cliniques fournit l'infrastructure nécessaire à l'utilisation secondaire des données, par exemple en recherche [97]. Un **lac de données** (*data lake*) conserve les données sous leur forme brute (tables, texte, images, fichiers de signal) et ne les structure qu'au moment de l'analyse. Le circuit des données dans un hôpital moderne part des systèmes sources et passe par des flux d'intégration jusqu'aux entrepôts et aux lacs de données, ce qui explique pourquoi il arrive qu'un clinicien ne trouve pas une information qui "existe" dans un autre système [98].

3.4.3 Le volume des données de santé

Le terme de mégadonnées (*big data*) ne renvoie pas seulement au volume. Les données de laboratoire, par exemple, présentent les caractéristiques typiques des mégadonnées : volume, vitesse, variété et véracité, et leur utilisation peut créer de la valeur dans la prise en charge des patients [99]. Ces cinq "V" peuvent être illustrés par quelques chiffres publiés :

- le jeu de séquences de génome entier de l'UK Biobank pour 500 000 participants occupe 1,10 tébioctet, même après conversion dans un format optimisé et compact [100]
- le réseau OHDSI comprend plus de 331 sources de données, avec 2,1 milliards de dossiers de patients provenant de 34 pays [101]
- un seul entrepôt hospitalo-universitaire peut contenir plus de 3 milliards d'observations cliniques [97].

Une cartographie systématique des architectures de mégadonnées en santé a désigné comme couches les plus souvent traitées le stockage, le traitement, l'ingestion, l'analyse et la visualisation. Les auteurs ont aussi signalé l'absence d'une couche automatique de vérification de la cohérence statistique [102].

3.5 Modèles communs de données pour la recherche

3.5.1 OMOP CDM et le réseau OHDSI

Chaque hôpital organise ses données à sa manière : d'autres tables, d'autres codes, d'autres unités. Une étude menée dans dix hôpitaux exigerait dix programmes d'analyse différents. Le modèle commun de données OMOP (*Observational Medical Outcomes Partnership Common Data Model*, OMOP CDM) standardise la structure et le contenu des données observationnelles, afin que la même analyse puisse être exécutée n'importe où [103]. La version en vigueur au 27 septembre 2026 est la 5.5 [103]. Les tables ont des noms fixes : PERSON pour les patients, VISIT_OCCURRENCE pour les visites, CONDITION_OCCURRENCE pour les diagnostics, DRUG_EXPOSURE pour les médicaments, MEASUREMENT pour les mesures.

Une structure commune ne suffit pas. Le contenu doit lui aussi utiliser les mêmes codes, et ce rôle revient aux vocabulaires standardisés OHDSI : plus de 10 millions de concepts issus de 136 vocabulaires, obligatoires pour toutes les sources du réseau [101]. La transformation des données locales au format OMOP se fait par un processus d'extraction, de transformation et de chargement (*extract, transform, load*, ETL) : les données sont extraites des systèmes sources, les codes locaux sont mis en correspondance avec les concepts standard, puis les données sont chargées dans les tables OMOP (figure 3.3). La conversion demande du temps et des ressources, c'est pourquoi des cadres ETL génériques, fondés sur des métadonnées, ont été développés [104].

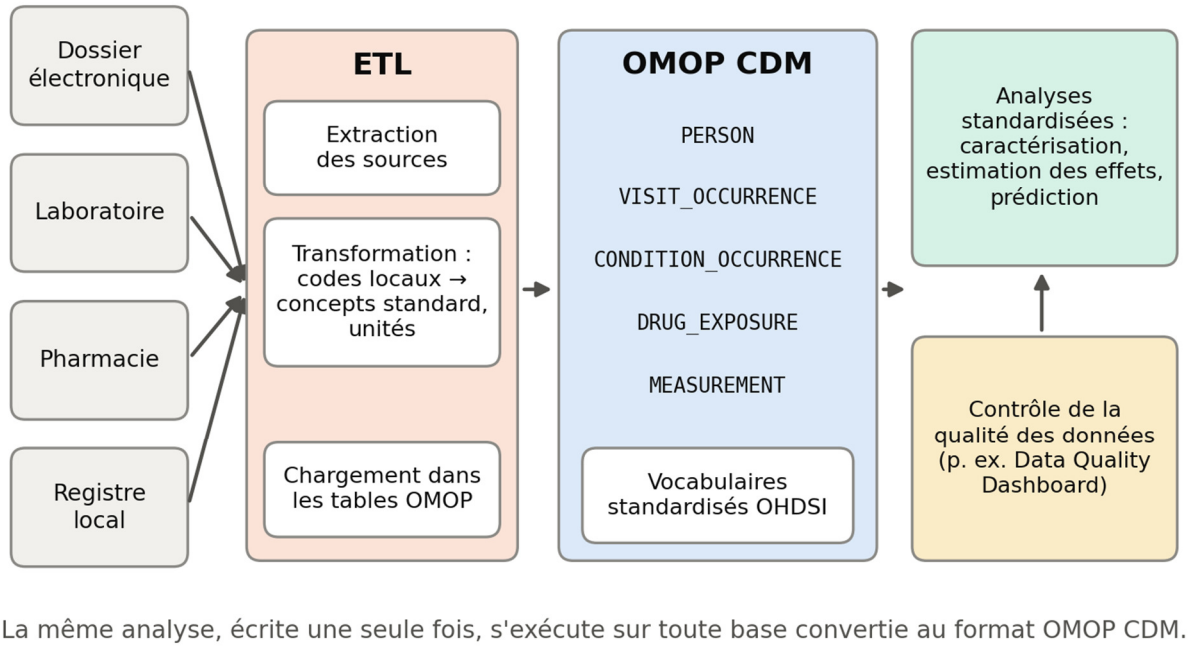


Figure 3.3. Des systèmes sources aux analyses standardisées : extraction, transformation (mise en correspondance avec les vocabulaires standard) et chargement dans l'OMOP CDM, suivis du contrôle de la qualité et d'analyses reproductibles. Schéma didactique d'après [101,103,105].

L'expérience européenne montre ce que suppose ce processus. Dans le projet EHDEN (*European Health Data & Evidence Network*), 25 partenaires de 11 pays ont converti des données concernant plus de 133 millions de patients. La durée médiane des 6 étapes clés de l'ETL a varié entre 4 et 115 jours. Les facteurs de réussite ont été une équipe adaptée, l'expertise dans la mise en correspondance des vocabulaires et la préparation précoce de la gouvernance des données [105]. La communauté OHDSI met l'accent sur la reproductibilité : protocoles, programmes et résultats publiés ouvertement, contrôles négatifs et positifs, analyse de plusieurs sources pour vérifier la généralisation. Les résultats ont inclus des études sur le traitement de l'hypertension et sur la COVID-19 [106]. Le nombre de publications consacrées à l'OMOP CDM a augmenté régulièrement après 2016 [107]. Une interrogation de PubMed réalisée pour cet ouvrage confirme la tendance et montre une croissance similaire pour les principes FAIR et pour REDCap (figure 3.4).

Un exemple simplifié montre ce que fait concrètement l'étape de transformation. Un hôpital fictif note dans son propre système le diagnostic "HTA essentielle" sous le code interne 4711, et le laboratoire rapporte la glycémie en mg/dL sous le code local GLU. Lors de l'ETL, le code 4711 est mis en correspondance avec le concept standard de l'hypertension essentielle, GLU avec le concept standard du glucose sérique, et l'unité de mesure est conservée explicitement. Après le chargement, une requête écrite une seule fois ("patients hypertendus dont la glycémie dépasse un seuil") produit des résultats comparables dans tous les hôpitaux du réseau. Dans EHDEN, la qualité de la conversion a également été vérifiée avec l'outil DataQualityDashboard [105].

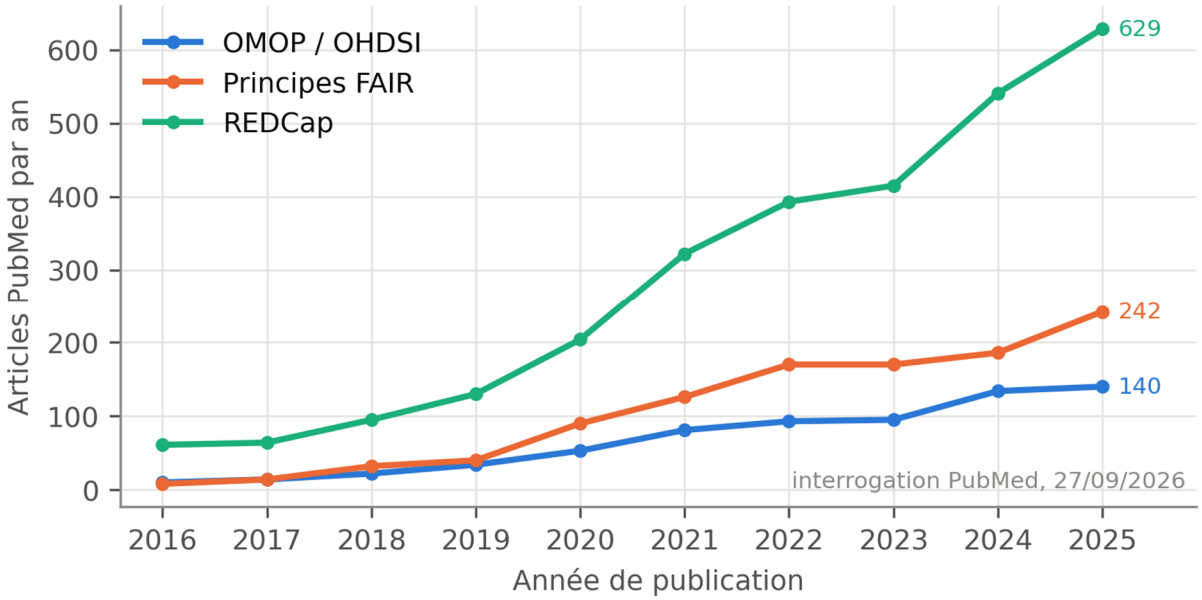


Figure 3.4. Nombre annuel d'articles indexés dans PubMed qui mentionnent dans le titre ou le résumé OMOP/OHDSI, les principes FAIR et REDCap, 2016-2025. Données réelles, interrogation E-utilities du 27 septembre 2026 [39] ; les valeurs reflètent les termes employés, et non le nombre total d'études qui utilisent ces outils.

3.5.2 i2b2

i2b2 (*Informatics for Integrating Biology and the Bedside*) est une plateforme libre (*open source*) d'interrogation et d'exploration des données cliniques, conçue surtout pour l'identification de cohortes et les études de faisabilité au moyen d'une interface graphique, sans programmation. La plateforme déclare plus de 250 hôpitaux utilisateurs [108]. Le modèle i2b2 repose classiquement sur une base relationnelle, avec une table centrale d'observations [97]. La différence pratique avec OMOP est une question d'accent. i2b2 répond rapidement à la question "combien de patients satisfont aux critères ?", alors que l'OMOP CDM, associé aux outils OHDSI, permet des analyses complètes et reproductibles en réseau.

3.5.3 Les principes FAIR en santé

Les principes FAIR exigent que les données soient faciles à trouver (*Findable*), accessibles (*Accessible*), interopérables (*Interoperable*) et réutilisables (*Reusable*). Concrètement, les données ont des identifiants uniques et pérennes et des métadonnées riches, elles sont enregistrées dans des ressources où l'on peut effectuer des recherches, elles peuvent être obtenues par un protocole standard et elles utilisent un langage formel pour la représentation des connaissances [109]. FAIR ne signifie pas "ouvert à tous". Les données de santé peuvent n'être accessibles qu'après authentification et autorisation, le principe exigeant seulement que le mode d'accès soit clair.

Pour les données de santé, le flux général de "FAIRisation" présente des lacunes liées à la curation, à la validation, à la désidentification, au versionnage et à l'indexation. Le projet européen FAIR4Health a proposé un flux révisé, avec une architecture fondée sur des ressources HL7 FHIR [110]. Une revue de portée (*scoping review*) de 34 articles a relevé, parmi les approches utilisées, des dictionnaires de données complets, l'interopérabilité sémantique, les ontologies et le contrôle de la qualité des données. Parmi les obstacles figuraient des coûts initiaux élevés, des problèmes techniques et administratifs et des préoccupations liées à la confidentialité [111].

3.6 La qualité des données et les registres médicaux

3.6.1 Dimensions et cadres d'évaluation

Le chapitre 2 a présenté les dimensions de la qualité : exactitude, complétude, actualité, cohérence. Les revues récentes confirment qu'il n'existe pas encore de standard unique d'évaluation de la qualité des données des DSE [55] et que, parmi ces dimensions, la cohérence influence toutes les autres [56]. Pour les études observationnelles, un cadre allemand propose 34 indicateurs regroupés en quatre aspects : l'intégrité (respect des exigences structurelles et techniques), la complétude (présence des valeurs), la cohérence (valeurs inadmissibles ou contradictoires) et l'exactitude (distributions et associations inattendues). Les indicateurs sont calculés automatiquement dans R, à partir des données et des métadonnées de l'étude [112]. Les auteurs recommandent de les utiliser aussi bien pour le suivi de la qualité en cours d'étude que pour l'analyse initiale des données, avant les analyses statistiques proprement dites [112].

En pratique, l'analyse initiale d'un fichier d'étude suit quelques étapes reproductibles. On compte les enregistrements et l'on recherche les doublons d'après l'identifiant. On calcule, pour chaque variable, la proportion de valeurs manquantes. On affiche le minimum et le maximum des variables numériques, afin de repérer les valeurs impossibles (une taille de 17 cm provient très probablement d'une faute de frappe pour la valeur 170). On vérifie ensuite les règles de cohérence entre champs. Chaque correction est consignée dans un journal, de sorte que le fichier final puisse être reconstitué à partir du fichier brut.

3.6.2 La validation à la saisie et le dictionnaire de données

L'erreur la moins coûteuse est celle qui n'entre pas dans la base de données. La validation à la saisie fait appel à quelques types de règles :

- **type et format** : la date a la forme AAAA-MM-JJ, le code numérique comporte le nombre correct de chiffres
- **intervalle** : pression systolique comprise entre 50 et 300 mmHg, âge compris entre 18 et 120 ans dans une étude chez l'adulte
- **liste de valeurs** : le sexe est F, M ou "non précisé", et non du texte libre
- **champ obligatoire** : la variable principale de l'étude ne peut pas manquer
- **cohérence entre champs** : la date de sortie ne précède pas la date d'admission, la pression diastolique est inférieure à la pression systolique.

Le **dictionnaire de données** (*data dictionary*, appelé aussi *codebook*) décrit chaque variable : le nom court, le libellé complet, le type, l'unité, les valeurs autorisées et le code utilisé pour les valeurs manquantes. Il constitue la part des métadonnées sans laquelle un fichier CSV ne peut pas être réutilisé par quelqu'un d'autre. C'est pourquoi il figure parmi les approches rapportées pour rendre les données de santé réutilisables [111]. Un exemple complet de dictionnaire de données est donné dans le TP 3.

3.6.3 Les registres cliniques

Un registre de patients recueille de manière systématique des données sur le diagnostic, les caractéristiques démographiques, le traitement et l'évolution d'un groupe défini de patients. Les registres peuvent être locaux, régionaux ou nationaux, généraux ou consacrés à une maladie [113]. Le programme SEER pour le cancer et le jeu de données de soins intensifs MIMIC-III, issu d'un seul hôpital, sont des exemples souvent étudiés. Parmi les limites citées figurent la faible implication des patients et le petit nombre de résultats rapportés par les patients [113]. La qualité

d'un registre se vérifie de façon systématique. Le programme de qualité des registres cardiovasculaires américains NCDR a fait état de données complètes et d'une exactitude très bonne, mais variable, avec une marge d'amélioration [114].

Les données hospitalières utilisées de façon rétrospective ont un caractère proche de celui d'un registre : dans un hôpital universitaire de Timişoara, les données de 867 patients opérés d'un cancer du côlon entre février 2019 et février 2023 ont été extraites et analysées pour comparer la période pandémique à la période antérieure [115], et une synthèse antérieure de la littérature sur le même thème a réuni les résultats rapportés par des centres de plusieurs pays, avec des baisses des interventions programmées comprises entre 14 % et 70 % [116].

3.7 Les outils de collecte électronique des données

3.7.1 Du formulaire papier à l'EDC

Le cahier d'observation (*case report form*, CRF) est l'outil qui sert à recueillir les données d'une étude. Dans sa version électronique (*electronic data capture*, EDC), il applique les règles de validation au moment même de la saisie et supprime la transcription à partir du papier. La ligne directrice de bonnes pratiques cliniques ICH E6(R3), adoptée le 6 janvier 2025, exige que les modifications des données sources soient traçables, sans masquer l'enregistrement initial, au moyen d'une piste d'audit (*audit trail*), et que les pistes d'audit ne puissent pas être désactivées [117]. Les données à caractère personnel dans l'Union européenne relèvent du règlement général sur la protection des données (RGPD). Les données concernant la santé sont des données à caractère personnel qui révèlent des informations sur l'état de santé, et la pseudonymisation signifie que les données ne peuvent plus être attribuées à une personne sans informations supplémentaires, conservées séparément [118].

3.7.2 REDCap

REDCap (*Research Electronic Data Capture*) a été créé à l'université Vanderbilt et lancé en 2004. Au 27 septembre 2026, la page de Vanderbilt fait état d'environ 3,8 millions d'utilisateurs, issus de près de 8 000 institutions de 163 pays, pour quelque 2,4 millions de projets [119]. Le logiciel est proposé gratuitement aux organisations à but non lucratif qui adhèrent au consortium, mais il n'est pas accessible aux personnes sans affiliation institutionnelle [120]. Il offre le contrôle des accès, une piste d'audit, des champs calculés, le marquage des champs qui identifient les participants et l'export en CSV ou dans les formats des logiciels SPSS, SAS, Stata et R [120]. REDCap a été présenté comme une solution possible aux difficultés de collecte, de gestion et de stockage des données, dans un contexte d'études de plus en plus complexes et de législation plus stricte [121]. Son adoption exige un soutien institutionnel : à l'université du Witwatersrand, le nombre de comptes actifs est passé de 129 en 2013 à 3 447 en 2021, avec des formations périodiques et une équipe locale d'assistance [122].

3.7.3 Google Forms et Epi Info

Google Forms est gratuit avec un compte Google et très facile à utiliser. Il permet de définir des règles de validation des réponses pour les questions à réponse courte (nombre compris dans un intervalle, longueur, expression régulière) [123] et de ramifier le questionnaire par sections en fonction de la réponse, pour les questions à choix multiple et à liste déroulante [124]. Les réponses peuvent être envoyées automatiquement dans une feuille Google Sheets, qui les organise en tableau [125]. Les plateformes en ligne comme Google Forms permettent de mener des

enquêtes et de recueillir les réponses d'un grand nombre de répondants, y compris en recherche médicale [126]. Elles ne sont toutefois pas conçues comme des systèmes EDC pour les essais cliniques réglementés.

Epi Info, développé par les CDC, était une suite de logiciels du domaine public destinée à la construction de formulaires et de bases de données, à la saisie des données et à l'analyse épidémiologique. Les CDC annoncent que le développement du produit et l'assistance technique ont été interrompus en septembre 2025, la version de bureau restant alors fonctionnelle sous Windows 11 [127]. Pour un nouveau projet, il faut peser cette situation : le logiciel peut être utilisé à des fins didactiques, mais sans mises à jour de sécurité et sans assistance.

Étude de cas

- Situation A : un groupe d'étudiants organise un questionnaire anonyme sur le sommeil, sans données d'identification, pour un projet de séminaire. Google Forms suffit, avec des règles de validation et l'export vers Google Sheets.
- Situation B : un service clinique prépare une étude multicentrique avec des données cliniques pseudonymisées, un suivi à 6 mois et un monitorage externe. Il faut un système EDC institutionnel, avec piste d'audit, rôles d'accès et export contrôlé, par exemple REDCap hébergé par l'université.
- Situation C : un cours d'épidémiologie appliquée utilise un ordinateur sans connexion Internet stable. Epi Info 7 peut servir pour l'exercice, à condition que les étudiants sachent que le produit n'est plus développé.

Tableau 3.3. Comparaison de trois outils de collecte électronique des données (situation au 27 septembre 2026). Source : [119,120,123,124,127,128].

Critère	REDCap	Google Forms	Epi Info 7
Coût	Gratuit pour les institutions à but non lucratif du consortium	Gratuit avec un compte Google	Domaine public, gratuit
Validation, ramification	Champs obligatoires, champs calculés	Règles par réponse, saut de sections	Module Check Code
Piste d'audit	Oui	Absente de la documentation consultée	Absente de la page des CDC
Conformité déclarée	RGPD, HIPAA, 21 CFR Part 11 (selon les auteurs)	Non conçu pour les études réglementées	Développement arrêté depuis septembre 2025
Export	CSV, SPSS, SAS, Stata, R	Google Sheets, CSV	Analyse intégrée au logiciel

Les affirmations de conformité du tableau concernent la plateforme. La conformité réelle d'une étude dépend de l'installation concrète, des procédures de l'institution et de la manière dont l'équipe utilise l'outil.

Pièges fréquents

- Utiliser le nom du patient comme clé d'identification : les noms se répètent, changent et s'écrivent de façons différentes.
- Mettre plusieurs valeurs dans une seule cellule ("I10, E11.9"), ce qui empêche le comptage et le filtrage.
- Conserver des fichiers CSV sans dictionnaire de données : un an plus tard, plus personne ne sait ce que signifie la colonne "tas2" ou le code 9.
- Confondre l'entrepôt de données (données nettoyées, structurées pour l'analyse) et le lac de données (données brutes).
- Croire que FAIR signifie données publiques. Les données de santé peuvent être FAIR tout en restant en accès contrôlé.
- Recueillir des identifiants (nom, code numérique personnel ou CNP, adresse électronique) dans un formulaire Google pour une étude qui n'en a pas besoin.
- Choisir un outil uniquement pour sa facilité d'utilisation, sans vérifier l'audit, l'export et l'assistance à long terme.

À retenir

- CSV est plat, JSON et XML sont hiérarchiques (semi-structurés). FHIR R5 utilise XML, JSON et Turtle.
- La clé primaire identifie une ligne de manière unique, et la clé étrangère la relie à une autre table.
- La normalisation (1NF-3NF) élimine la redondance : chaque fait n'est écrit qu'une seule fois.
- SELECT, WHERE, JOIN et GROUP BY couvrent la plupart des requêtes cliniques simples.
- ACID : atomicité, cohérence, isolation, durabilité.
- L'OMOP CDM standardise la structure, et les vocabulaires OHDSI standardisent le contenu des données observationnelles.
- FAIR : facile à trouver, accessible, interopérable, réutilisable, pas nécessairement ouvert.
- L'outil EDC se choisit en fonction de la validation, de la piste d'audit, de l'export et de la conformité, et pas seulement du coût.

Questions d'autoévaluation

1. (QCS) Dans une base de données relationnelle, la colonne `id_pacient` de la table Consultation, qui désigne une ligne de la table Patient, est :

a) la clé primaire de la table Consultation
b) une clé étrangère
c) un index de texte libre
d) une métadonnée du fichier
e) une clé de chiffrement

2. (QCS) Une table dans laquelle la cellule "diagnostics" contient la valeur "I10, E11.9" enfreint :

a) la première forme normale
b) la propriété de durabilité
c) le principe FAIR "facile à trouver"
d) le format JSON
e) l'isolation des transactions

3. (QCS) Le rôle principal des vocabulaires standardisés OHDSI est :

a) le chiffrement des données de santé
b) la standardisation du contenu (des codes) des données du réseau
c) le remplacement du langage SQL
d) le stockage des images DICOM
e) la collecte des données au moyen de formulaires web

4. (QCM) À propos des propriétés ACID des transactions, quelles affirmations sont exactes ?

a) l'atomicité signifie que toutes les opérations de la transaction s'achèvent, ou aucune
b) la durabilité signifie qu'une modification validée résiste à une panne du système
c) l'isolation signifie que les données sont chiffrées
d) la cohérence signifie le respect des contraintes d'intégrité
e) ACID est un format de fichier pour les données génomiques

5. (QCM) À propos des outils de collecte électronique des données, quelles affirmations sont exactes ?

a) REDCap offre une piste d'audit et l'export vers des logiciels statistiques
b) Google Forms permet de ramifier le questionnaire par sections en fonction de la réponse
c) ICH E6(R3) exige que les modifications des données soient traçables grâce à une piste d'audit
d) le développement d'Epi Info a été interrompu par les CDC en septembre 2025
e) Google Forms est conçu comme un système EDC pour les essais cliniques réglementés

CHAPITRE 4

Chapitre 4. Classifications, nomenclatures et terminologies médicales

Objectifs

À la fin de ce chapitre, l'étudiant sera capable :

- de définir et de distinguer, à partir d'exemples, une classification, une nomenclature, une terminologie clinique et une ontologie
- de décomposer un code CIM-10 en chapitre, bloc, catégorie et sous-catégorie et d'expliquer pourquoi le codage part de l'index, mais se vérifie dans la liste tabulaire
- d'appliquer la définition de la cause initiale du décès à un certificat de décès simple
- de construire un cluster CIM-11 à partir d'un code souche et de codes d'extension, en utilisant correctement les séparateurs "&" et "/"
- de comparer la CIM-10, la CIM-11, SNOMED CT, LOINC, l'ATC et MeSH selon leur finalité, leur propriétaire et leur mode d'utilisation
- de décrire le circuit du groupage DRG en Roumanie et de reconnaître les risques liés à la qualité du codage et au codage automatique.

4.1 Pourquoi coder

4.1.1 Quatre raisons pratiques

Un médecin écrit "infarctus antérieur étendu, STEMI", son collègue d'un autre hôpital écrit "IDM avec sus-décalage du segment ST dans le territoire de l'IVA". Pour un ordinateur qui compte les cas d'un pays, ce sont deux chaînes de caractères différentes. Le codage transforme la formulation libre en un symbole standardisé, qui peut être compté, comparé et transmis.

La première raison est la **statistique de mortalité**. Depuis son adoption, la Classification internationale des maladies (CIM, en anglais *International Classification of Diseases*, ICD) constitue la base principale de la comparaison des statistiques des causes de décès et de maladie entre pays et dans le temps [129]. La CIM a une longue histoire : la première édition est parue en 1893, et les révisions se sont succédé à peu près tous les dix ans, d'abord sous l'égide de l'Institut international de statistique, puis sous celle de l'Organisation mondiale de la santé (OMS) [15].

La deuxième raison est la **statistique de morbidité** : combien d'hospitalisations pour pneumonie un département a-t-il enregistrées, quelle proportion des patients hospitalisés a un diabète. Ce qui compte ici, c'est le diagnostic principal de l'épisode de soins, notion définie plus loin (4.2.3).

La troisième raison est le **financement**. Dans les systèmes fondés sur les groupes homogènes de malades (*diagnosis-related groups*, DRG), les codes de diagnostic et d'acte déterminent le groupe dans lequel le cas est classé et, par ce biais, le montant remboursé (4.6). En Australie, par exemple, l'ICD-10-AM sert de base au financement à l'activité, et les auteurs d'une analyse de la littérature observent que cette fonction a réorienté les objectifs du codage clinique [130].

La quatrième raison est la **recherche et l'interopérabilité sémantique**. Les données codées sont réutilisées dans des études épidémiologiques : une revue de portée (*scoping review*) a recensé 636 articles publiés entre 2012 et 2022 qui utilisaient des données australiennes codées en CIM-10 ou en ICD-10-AM, pour la plupart des études observationnelles [130]. L'interopérabilité sémantique signifie que le système qui reçoit les données en comprend le sens exact dans lequel elles ont été envoyées. Les standards de messagerie transportent la structure, mais ce sont les codes des terminologies qui donnent le sens.

4.1.2 Classification, nomenclature, terminologie, ontologie

Les termes terminologie, vocabulaire et nomenclature sont souvent employés comme synonymes, y compris par les concepteurs des systèmes de codage [15]. Les différences entre les quatre notions présentées ci-dessous comptent lorsqu'il s'agit de choisir l'outil.

Une **classification** est un ensemble de classes qui regroupent des concepts que l'on ne souhaite plus distinguer dans l'analyse. Une classification des diagnostics, comme la CIM-11, se compose de classes de diagnostics, et une classification des interventions, comme l'ICHI, de classes d'interventions [131]. Les classes sont exhaustives (tout cas aboutit quelque part, éventuellement dans une classe "autres" ou "sans précision") et mutuellement exclusives, ce qui permet de compter sans doublons.

Une **nomenclature** fournit des termes assortis de codes qui servent à étiqueter (indexer) les objets de la médecine : cas, patients, documents. SNOMED CT est l'exemple donné par le manuel de Winter et collaborateurs [131]. La nomenclature ne regroupe pas. Elle cherche à nommer chaque concept clinique aussi précisément que possible.

Par **terminologie clinique**, on entend, dans la pratique de l'informatique médicale, une nomenclature contrôlée, dotée d'identifiants stables, de synonymes et de relations entre concepts. Toute terminologie standard a besoin d'un processus de maintenance continu, faute de quoi elle vieillit vite [15].

Une **ontologie** ajoute des définitions logiques. Les concepts sont reliés par des relations formelles ("est un", "a pour localisation", "est causé par"), et un programme peut déduire des relations qui n'ont pas été écrites explicitement. Les ontologies permettent la représentation structurée et l'intégration des connaissances médicales et servent à l'interopérabilité et à l'aide à la décision, mais elles sont difficiles à standardiser et à mettre à jour [132].

Une comparaison simple aide à y voir clair. Dans la CIM-10, la pneumopathie bactérienne sans précision reçoit la sous-catégorie J15.9, une classe qui sert au dénombrement [133]. Dans SNOMED CT, le concept "pneumonie bactérienne" est défini par des relations : c'est une maladie infectieuse du poumon, qui a pour localisation "structure du poumon", pour morphologie "inflammation" et pour agent causal "bactérie" [15]. La première construction répond à la question "combien ?", la seconde à la question "quoi exactement ?".

4.2 La CIM-10 et les modifications nationales

4.2.1 Structure

La CIM-10 a été publiée par l'OMS en 1992 [15]. La version OMS actuelle est celle de 2019, disponible dans le navigateur officiel et accompagnée du volume 2 (manuel d'utilisation, 6e édition) [133]. La classification comporte trois volumes : le volume 1 contient la liste tabulaire, le volume 2 donne les règles d'utilisation et le volume 3 est l'index alphabétique [133].

Le noyau est une liste de **catégories à trois caractères**, chacune pouvant être subdivisée en dix **sous-catégories à quatre caractères** au plus. Le code commence par une lettre suivie de chiffres, et le quatrième caractère vient après le point, par exemple I21.0 [133]. Le niveau à trois caractères est obligatoire pour la transmission des données à la base de mortalité de l'OMS et pour les comparaisons internationales [133]. Les catégories sont regroupées en blocs (intervalles, comme I20-I25), et les blocs en chapitres numérotés en chiffres romains. La version 2019 compte 22 chapitres, le dernier étant le chapitre XXII, Codes d'utilisation particulière (U00-U85), où se trouvent aussi les codes de la COVID-19 [133]. La figure 4.1 suit un code depuis le chapitre jusqu'à la sous-catégorie.

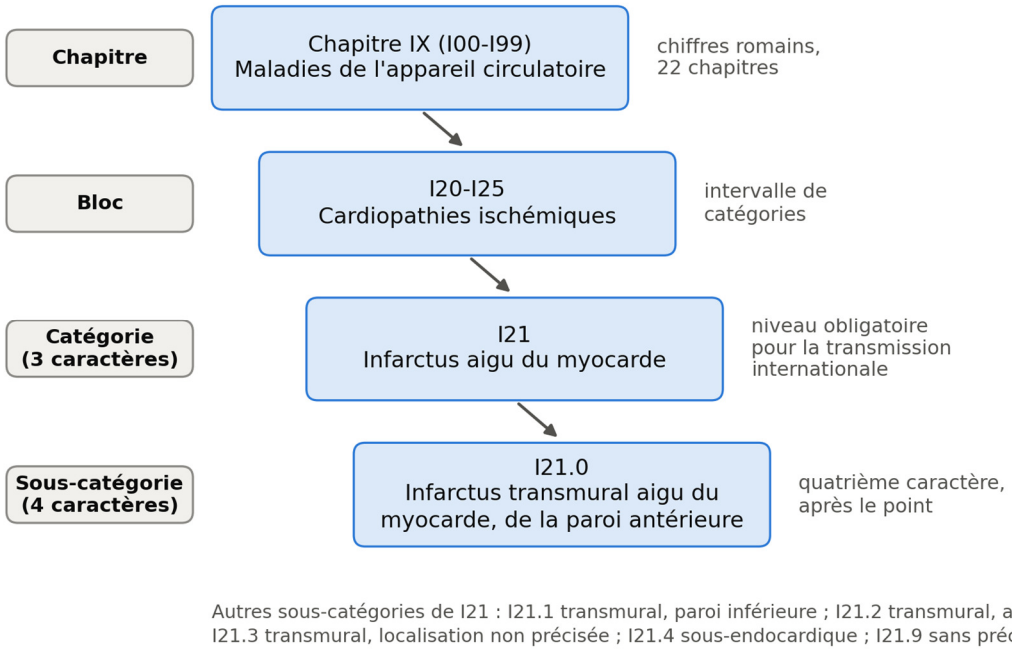


Figure 4.1. Structure hiérarchique de la CIM-10, à partir de l'exemple de l'infarctus transmural aigu du myocarde, de la paroi antérieure (I21.0). Codes et libellés vérifiés dans le navigateur CIM-10 de l'OMS, version 2019 [133].

Deux conventions reviennent souvent dans les exercices. La terminaison .8 signifie généralement "autres formes", et .9 "sans précision" [15,133]. Le système "dague et astérisque" attribue deux codes à un diagnostic qui comporte à la fois la maladie générale et sa manifestation dans un organe : le code de la maladie sous-jacente reçoit la dague (†), et le code facultatif de la manifestation reçoit l'astérisque (*) [133]. La polyneuropathie diabétique chez un patient atteint de diabète de type 2 s'écrit ainsi E11.4† G63.2*.

4.2.2 Règles de base : l'index, puis la liste tabulaire

L'index alphabétique contient de nombreux termes qui ne figurent pas dans le volume 1. C'est pourquoi le codage exige de consulter les deux parties avant d'attribuer le code [133]. Les étapes recommandées par l'OMS pour l'utilisateur occasionnel sont, en résumé [133] :

1. identifier le type d'énoncé (maladie ou lésion traumatique, ou bien cause externe) et la section correspondante de l'index
2. rechercher le terme principal (*lead term*), en général le substantif qui désigne l'état pathologique
3. lire les notes, les modificateurs entre parenthèses (ils ne changent pas le code) et les termes en retrait (ils peuvent changer le code)

4. suivre les renvois "voir" et "voir aussi"
5. vérifier dans la liste tabulaire le code trouvé, y compris les notes d'inclusion et d'exclusion
6. attribuer le code.

Un code de l'index suivi d'un tiret (par exemple "I21.-") signale qu'il existe un quatrième caractère à rechercher dans la liste tabulaire [133].

4.2.3 Le diagnostic principal et la cause initiale du décès

Pour l'analyse de morbidité fondée sur une affection unique, l'OMS définit l'**affection principale** comme l'affection diagnostiquée à la fin de l'épisode de soins, principalement responsable du besoin de traitement ou d'examens du patient. S'il y en a plusieurs, on retient celle qui a mobilisé le plus de ressources. En l'absence de diagnostic établi, on retient le symptôme principal [133].

Pour la mortalité, on tabule la **cause initiale du décès**, définie comme "(a) la maladie ou le traumatisme qui a déclenché l'évolution morbide conduisant directement au décès, ou (b) les circonstances de l'accident ou de la violence qui ont entraîné le traumatisme mortel" [133]. Cette définition répond à une logique de santé publique. La prévention la plus efficace interrompt la chaîne à son début [133].

Un exemple tiré du guide de l'OMS montre la différence entre la cause terminale et la cause initiale. Le certificat indique : (a) pneumonie, due à (b) une fracture de la hanche, due à (c) un trébuchement sur un tapis. Le point de départ de la séquence est le trébuchement, c'est-à-dire l'événement externe [134]. La statistique ne comptera pas ce décès sous "pneumonie". La sélection se fait en deux étapes : on identifie le point de départ, puis on vérifie si les règles particulières le modifient [133].

4.2.4 Les modifications nationales et l'ICD-10-AM

L'OMS a arrêté la maintenance de la CIM-10 en 2018, mais les pays peuvent l'utiliser aussi longtemps qu'ils en ont besoin [135]. De nombreux États ont élaboré leurs propres modifications cliniques : les États-Unis ont l'ICD-10-CM, le Canada la CIM-10-CA, la Corée la KCD, l'Australie l'ICD-10-AM [130,136-138]. Pour l'ICD-10-CM, l'écart de granularité est important. Fung et collaborateurs ont dénombré 10 607 codes terminaux dans la CIM-10 (OMS), contre 71 932 dans l'ICD-10-CM [136]. L'Australie utilise la CIM-10 pour la mortalité et l'ICD-10-AM pour la morbidité [130].

La situation en Roumanie en 2026

- Les diagnostics des patients hospitalisés sont codés en ICD-10-AM, 3e édition (Australie, juillet 2002), et les actes selon la "Clasificarea RO.vi.DRG a procedurilor, ed. I" (classification RO.vi.DRG des actes, 1re édition), dérivée de la liste australienne ACHI [139].
- La classification DRG en vigueur est la RO DRG v.1.1 (mars 2024), qui compte 669 groupes : 651 dans les 23 catégories majeures de diagnostic, 12 groupes pré-CMD et 6 groupes d'erreur [139].
- La classification est maintenue par l'Institut national de management des services de santé (Institutul Național de Management al Serviciilor de Sănătate, INMSS). Son histoire commence en 2001, et elle est utilisée pour le financement DRG depuis 2003 [139].
- Au 27 septembre 2026, les groupages DRG en Roumanie reposent encore sur l'ICD-10-AM, 3e éd. Nous n'avons identifié ni acte normatif ni calendrier officiel pour l'adoption de la CIM-11.

4.3 La CIM-11

4.3.1 Adoption et situation en 2026

La CIM-11 a été adoptée par l'Assemblée mondiale de la santé en mai 2019 [129] et est entrée en vigueur le 1er janvier 2022 [140]. La version actuelle du navigateur officiel est la CIM-11 MMS, version 2026-01 [141]. L'aide-mémoire de l'OMS de mai 2026 fait état de 132 pays à différents stades de mise en œuvre et de plus de 17 000 catégories de diagnostic [142]. Une ventilation de mai 2024, encore affichée par l'OMS, montrait que seuls 14 pays et territoires parmi eux collectaient ou transmettaient déjà des données en CIM-11 [140].

4.3.2 Architecture : la Fondation et les linéarisations

Le changement majeur par rapport à la CIM-10 ne tient pas au nombre de codes, mais à la construction. La CIM-11 comprend une base de connaissances sémantique, la **Fondation** (*Foundation Component*), une ontologie biomédicale qui lui est liée et des classifications dérivées de la Fondation. La principale classification dérivée, qui succède à la CIM-10, est la **CIM-11 pour les statistiques de mortalité et de morbidité** (MMS) [129]. Le processus par lequel les entités de la Fondation sont extraites dans une classification destinée à un usage particulier s'appelle **linéarisation**. Le contenu et le sens de l'entité ne changent pas lors de l'extraction [134]. À partir de la même Fondation, on peut générer des linéarisations de spécialité et des modifications nationales, ces dernières uniquement en accord avec l'OMS [134]. La figure 4.2 résume cette architecture.

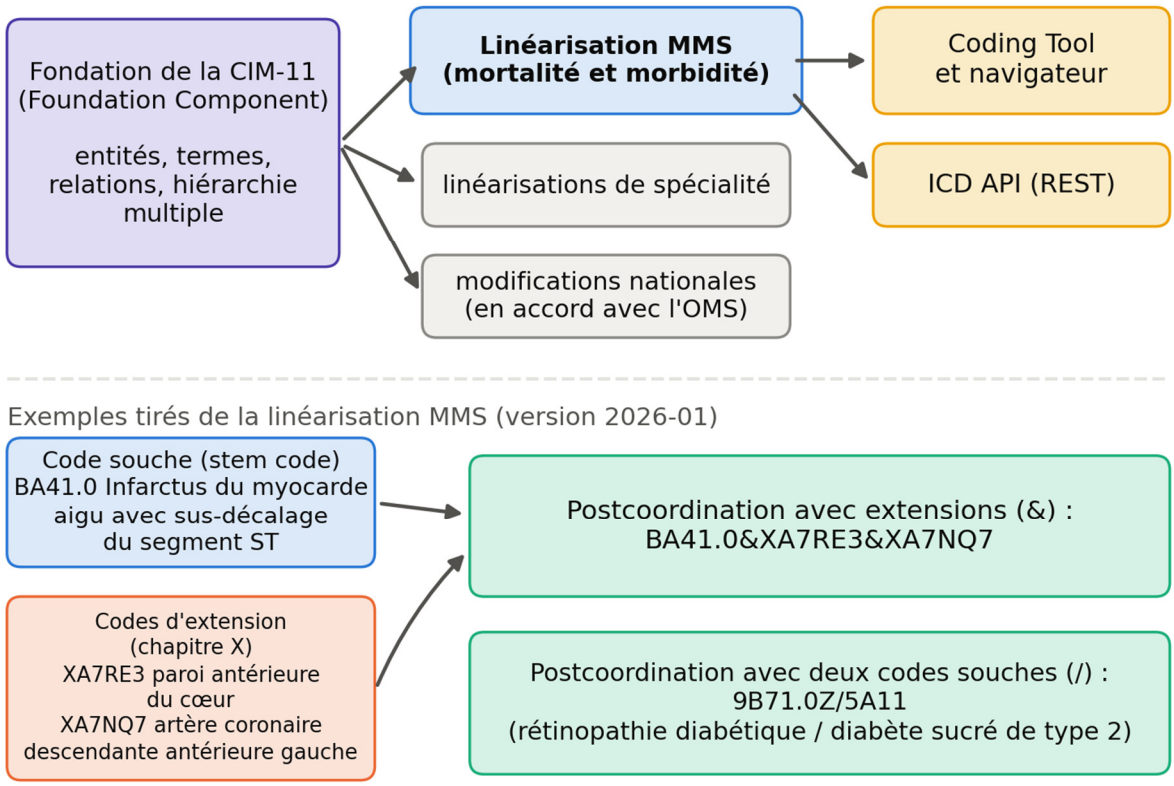


Figure 4.2. Architecture de la CIM-11 : la Fondation, les linéarisations et les outils de l'OMS, avec des exemples de codes souches, de codes d'extension et de clusters de la linéarisation MMS, version 2026-01. Synthèse d'après [129,134]. Les exemples de clusters sont repris du guide de référence de l'OMS [134].

4.3.3 Le code CIM-11 : structure, codes souches et codes d'extension

Les codes MMS sont alphanumériques et couvrent l'intervalle 1A00.00-ZZ9Z.ZZ. Le premier caractère indique toujours le chapitre, le deuxième est toujours une lettre (pour éviter toute confusion avec un code CIM-10), et les lettres O et I sont absentes, afin de ne pas être confondues avec les chiffres 0 et 1 [134]. La terminaison Y est réservée à la catégorie "autres précisés", et la terminaison Z à la catégorie "sans précision" [134]. Ainsi, BA41.Z signifie infarctus aigu du myocarde, sans précision.

Les codes qui peuvent être utilisés seuls s'appellent **codes souches** (*stem codes*) [134]. Les **codes d'extension** commencent par la lettre X, forment le chapitre X et ajoutent des détails : anatomie, agent infectieux, histopathologie, sévérité, médicament [134]. Ils ne s'utilisent pas seuls en statistique et ne peuvent jamais occuper la première position d'un cluster [134].

4.3.4 La postcoordination

La **postcoordination** relie, lorsque c'est nécessaire, deux codes souches ou un code souche et un ou plusieurs codes d'extension. Le groupe obtenu s'appelle **cluster** et décrit un seul concept clinique [134]. La syntaxe est fixe [134] :

- "&" relie un code souche à ses codes d'extension : code souche&extension1&extension2
- "/" relie deux codes souches : code1/code2
- lorsqu'un cluster contient deux codes souches, chacun avec ses extensions, les extensions s'écrivent immédiatement après le code souche auquel elles appartiennent.

L'inverse est la **précoordination** : un seul code souche contient déjà tous les éléments. Le carcinome épidermoïde des bronches ou du poumon, 2C25.2, réunit la localisation et l'histologie en un seul code [134].

Exemple

Diagnostic (cas fictif) : infarctus aigu du myocarde avec sus-décalage du segment ST, paroi antérieure, lésion de l'artère coronaire descendante antérieure gauche.

- code souche : BA41.0 Infarctus du myocarde aigu avec sus-décalage du segment ST
- extension d'anatomie : XA7RE3 Paroi antérieure du cœur
- extension d'anatomie : XA7NQ7 Artère coronaire descendante antérieure gauche
- cluster : **BA41.0&XA7RE3&XA7NQ7**

Diagnostic (cas fictif) : fracture sous-capitale du col du fémur gauche, fermée, après un trébuchement et une chute de plain-pied.

- cluster : **NC72.20&XK8G&XJ44E/PA60** (la fracture, la latéralité gauche, la fracture fermée, puis la cause externe, chute accidentelle de plain-pied ou d'une hauteur de 1 mètre ou moins)

Les codes ont été vérifiés dans le navigateur CIM-11 MMS, version 2026-01 [141]. Le premier exemple reproduit un exemple du guide de référence de l'OMS [134]. En CIM-10, le premier cas reçoit I21.0, sans l'information sur l'artère atteinte, et le second S72.0, avec W01 comme code de cause externe [133].

L'ordre des codes souches dépend de la finalité. En morbidité, le premier code est l'affection principale. En mortalité, le premier code est la cause initiale. Le guide de l'OMS donne l'exemple du diabète de type 2 compliqué d'un coma diabétique : à l'admission, le cluster est 5A23/5A11 (le coma est le motif de l'hospitalisation), alors que sur le certificat de décès il est 5A11/5A23 (le diabète a déclenché la chaîne) [134].

4.3.5 Les outils de l'OMS : navigateur, Coding Tool, API

Dans la CIM-11, l'index imprimé est remplacé par l'outil de codage en ligne (*Coding Tool*), et l'interface de programmation d'applications (*application programming interface*, API) permet d'accéder à distance au contenu et aux services de la CIM-11 [129]. Le Coding Tool effectue la recherche par termes et peut renvoyer directement un cluster assemblé, et pas seulement un code isolé [134]. Grâce à l'API, un hôpital peut intégrer la recherche CIM-11 directement dans le formulaire de sortie du dossier de santé électronique.

4.3.6 Nouveaux chapitres

La CIM-11 compte cinq nouveaux chapitres. Les maladies du sang (chapitre 03) et celles du système immunitaire (chapitre 04) sont séparées. Les troubles du cycle veille-sommeil ont leur propre chapitre (07). Les affections liées à la santé sexuelle sont regroupées dans le chapitre 17. La médecine traditionnelle dispose du chapitre supplémentaire 26 [134]. S'y ajoutent une section supplémentaire pour l'évaluation du fonctionnement (V), alignée sur la CIF, et le chapitre X des codes d'extension [134].

Tableau 4.1. Comparaison de la CIM-10 et de la CIM-11. Source : [129,133-136].

Caractéristique	**CIM-10**	**CIM-11 (MMS)**
Approbation / entrée en vigueur	publiée en 1992 ; version OMS finale 2019	adoptée en 2019 ; en vigueur depuis le 1er janv. 2022
Format du code	lettre + chiffres, 3-4 caractères (A00.0-Z99.9)	4-6 caractères, 1A00.00-ZZ9Z.ZZ ; deuxième position toujours une lettre
Codes terminaux (sans extensions)	10 607	14 622 (hors chapitres 26, V et X)
Chapitres	22 (version 2019)	26 + section V + chapitre X
Détails supplémentaires	dague/astérisque, codes multiples	postcoordination avec "&" et "/"
Recherche	index alphabétique (volume 3)	Coding Tool, API
Maintenance	arrêtée depuis 2018	mises à jour annuelles

4.3.7 La transition : ce que montrent les évaluations

L'essai sur le terrain mené à Calgary (Canada) a recodé dans la version bêta 2018 de la CIM-11 des dossiers d'hospitalisation codés auparavant en CIM-10-CA. Le code de l'affection principale a concordé dans 86,3 % des cas, et le temps moyen de codage est passé de 23,6 à 9,9 minutes par dossier à mesure que les codeurs gagnaient en expérience [137]. En Corée, 27 spécialistes de la gestion de l'information de santé ont obtenu, pour le codage de termes diagnostiques isolés, une exactitude moyenne de 71,6 % en CIM-11, contre 80,2 % dans la classification nationale dérivée de la CIM-10. Les clusters, les termes modifiés et les codes supprimés ont contribué à la moindre concordance [138].

Le contenu de la CIM-11 couvre bien les anciennes classifications si l'on recourt à la postcoordination. Sur 1 725 codes ICD-10-CM, la couverture complète atteignait 35,2 % avec les seuls codes souches, 46,5 % après l'ajout des entités de la Fondation et 89,4 % avec la postcoordination (figure 4.3) [143].

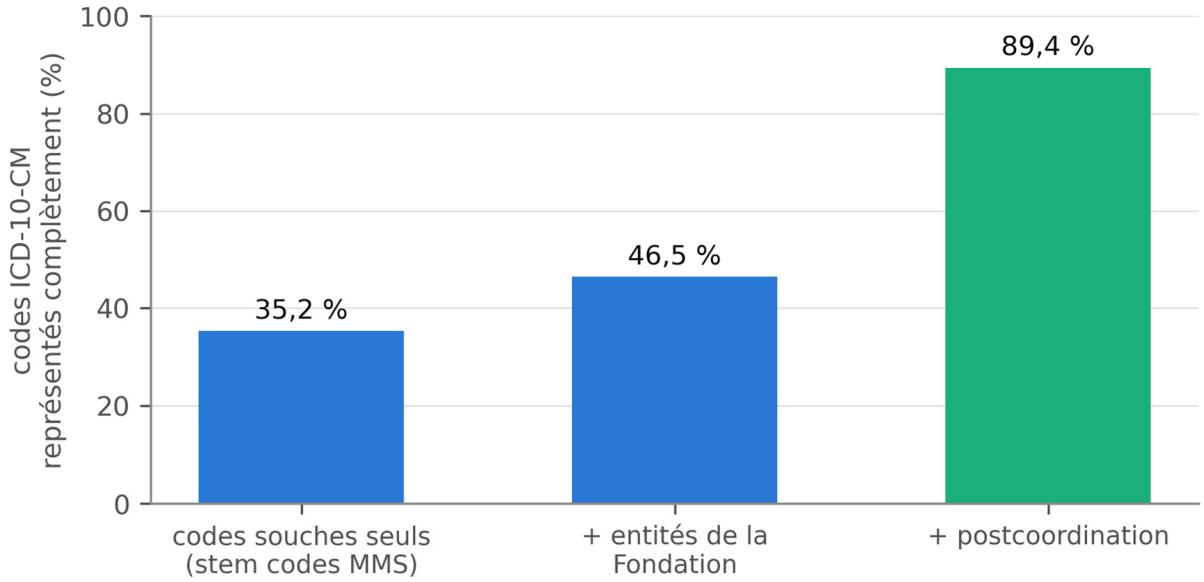


Figure 4.3. Proportion des codes ICD-10-CM (n = 1 725) représentés complètement dans la CIM-11, en cumul, par étapes : codes souches MMS, entités de la Fondation, postcoordination. Données réelles d'après [143].

Les tables de correspondance (*mapping tables*) CIM-10-CIM-11 se téléchargent depuis l'onglet "Info" du navigateur. L'OMS avertit qu'elles servent à comparer les données entre révisions, et non à la conversion automatique [135]. Appliquées à 24 645 décès survenus en Alberta, les tables ont permis de mettre en correspondance 96,6 % des codes CIM-10 des causes de décès chez l'adulte [144]. En revanche, dans la correspondance canadienne de la CIM-11 vers la CIM-10-CA, 65 % des codes CIM-10-CA cibles étaient plus larges que le code CIM-11 source, et 26 % seulement étaient équivalents [145]. Autrement dit, le retour vers la CIM-10 fait perdre du détail.

4.4 Autres classifications de l'OMS

La **Classification internationale du fonctionnement, du handicap et de la santé** (CIF) a été approuvée par l'Assemblée mondiale de la santé en 2001, par la résolution WHA54.21. Elle décrit les fonctions organiques et les structures anatomiques, les activités et la participation, ainsi que les facteurs environnementaux, dans un modèle biopsychosocial [146]. La CIM dit de quelle maladie souffre le patient, la CIF dit ce qu'il peut faire avec cette maladie.

La **Classification internationale des interventions de santé** (ICHI) décrit les interventions selon trois axes : la cible (l'entité sur laquelle on agit), l'action et les moyens. La version bêta-3 a été publiée en octobre 2020, et la composante consacrée aux interventions de santé publique a été finalisée en 2023 [147]. La CIM, la CIF et l'ICHI utilisent le même ensemble de codes d'extension [146].

4.5 Terminologies cliniques

4.5.1 SNOMED CT

SNOMED International présente SNOMED CT comme la terminologie clinique multilingue la plus complète au monde. L'édition internationale dépasse 360 000 concepts, fait l'objet d'une publication mensuelle et est utilisée dans plus de 80 pays [148]. Elle comporte trois composantes : les **concepts** (chacun doté d'un identifiant numérique unique), les **descriptions** (les termes lisibles par l'humain, y compris les synonymes) et les **relations** entre concepts [148]. La relation "est un"

(*is a*) construit les hiérarchies à partir d'un concept racine, et des attributs comme "localisation" (*finding site*) et "morphologie associée" définissent logiquement le concept [149]. Un concept peut avoir plusieurs parents. La "pneumonie bactérienne" est, par exemple, à la fois une maladie infectieuse du poumon et une inflammation des voies respiratoires inférieures [15].

SNOMED CT permet elle aussi la postcoordination, au moyen d'**expressions** écrites dans une grammaire compositionnelle. L'expression "80146002 |appendectomy| : 260870009 |priority| = 25876001 |emergency|" décrit une appendicectomie en urgence [149]. Cette liberté a un prix. La même "appendicite aiguë" peut être codée comme un seul concept, comme un modificateur plus une maladie ou comme un modificateur plus une morphologie plus une topographie, et deux codeurs peuvent aboutir à des expressions différentes pour le même sens [15].

Une revue de 651 articles publiés entre 2020 et 2025 montre que SNOMED CT est passée de la phase théorique à la mise en œuvre. Les études de mise en œuvre et d'évaluation représentaient respectivement 26,4 % et 32,0 % du total, et l'utilisation la plus fréquente était l'extraction de données de patients à partir de registres [150]. Les mêmes auteurs notent que l'uniformité du codage reste un problème et que les preuves d'une amélioration des résultats pour les patients font défaut [150]. SNOMED International compte 54 membres nationaux et a délivré des licences d'affilié à plus de 50 000 personnes et organisations. La Roumanie ne figure pas parmi les membres au 27 septembre 2026 [151].

4.5.2 LOINC

LOINC (*Logical Observation Identifiers Names and Codes*) nomme et code les analyses de laboratoire et d'autres observations cliniques. Le nom complet d'un terme LOINC comporte cinq ou six parties, séparées par ":" [152] :

1. le **composant** (l'analyte : glucose, créatinine)
2. la **propriété** (concentration de masse, concentration molaire, nombre)
3. l'**aspect temporel** (instant ou intervalle, par exemple les urines de 24 heures)
4. le **système** (le type d'échantillon : sérum, plasma, urine)
5. l'**échelle** (quantitative, ordinale, nominale)
6. la **méthode**, uniquement lorsqu'elle compte pour l'interprétation.

La glycémie sérique ou plasmatique a le code 2345-7, avec le nom complet Glucose:MCnc:Pt:Ser/Plas:Qn [153]. Le code numérique n'a pas de structure interne, à l'exception du dernier chiffre, qui est un chiffre de contrôle [152].

4.5.3 ATC/DDD pour les médicaments

La classification anatomique, thérapeutique et chimique (ATC) est maintenue par le Centre collaborateur de l'OMS pour la méthodologie des statistiques sur les médicaments, hébergé par l'Institut norvégien de santé publique. Elle comporte cinq niveaux et 14 groupes anatomiques principaux [154]. La metformine, A10BA02, se lit ainsi : A voies digestives et métabolisme, A10 médicaments du diabète, A10B hypoglycémiants, insulines exclues, A10BA biguanides, A10BA02 metformine [154]. La dose définie journalière (*defined daily dose*, DDD) est la dose d'entretien moyenne supposée par jour pour l'indication principale chez l'adulte. C'est une unité de mesure destinée aux études de consommation, qui ne reflète pas nécessairement la dose recommandée ou prescrite [155].

4.5.4 MeSH et UMLS

MeSH (*Medical Subject Headings*) est le vocabulaire contrôlé et hiérarchisé de la Bibliothèque nationale de médecine des États-Unis (NLM), utilisé pour l'indexation, le catalogage et la recherche de l'information biomédicale et mis à jour chaque année [156]. Un terme peut apparaître à plusieurs endroits de la hiérarchie [15]. MeSH ne sert pas au codage des patients. C'est l'outil avec lequel on recherche la littérature dans PubMed (chapitre 1).

L'UMLS (*Unified Medical Language System*) réunit des vocabulaires et des standards, parmi lesquels l'ICD-10-CM, LOINC, MeSH, RxNorm et SNOMED CT. Il comporte trois sources de connaissances : le Métathésaurus, le Réseau sémantique et le Lexique SPECIALIST [157].

Tableau 4.2. Principaux systèmes de classification et terminologies. Source : [140,146-148,152,154,156,157].

Système	Type	Propriétaire	Domaine	Utilisation principale
CIM-10 / CIM-11	classification	OMS	maladies, causes de décès	statistique, financement
CIF	classification	OMS	fonctionnement, handicap	réadaptation, politiques sociales
ICHI	classification	OMS	interventions	déclaration des interventions
SNOMED CT	terminologie clinique	SNOMED International	concepts cliniques	dossier électronique
LOINC	terminologie	Regenstrief Institute	analyses, observations	résultats de laboratoire
ATC/DDD	classification	Centre collaborateur de l'OMS (Oslo)	substances actives	statistiques de consommation
MeSH	thésaurus	NLM	sujets biomédicaux	recherche documentaire
UMLS	métathésaurus	NLM	vocabulaires multiples	correspondances entre vocabulaires

4.6 Les groupes homogènes de malades (DRG)

4.6.1 Le principe

Le DRG est un système de classification des patients qui relie le nombre et les types de patients traités dans un hôpital aux ressources consommées [139]. Chaque groupe a un **poids relatif** (*relative weight*), qui exprime la consommation de ressources attendue pour ce type d'hospitalisation par rapport à l'hospitalisation moyenne [158]. La moyenne des poids des cas sortis d'un hôpital constitue l'**indice de case-mix** (ICM) [159]. Un ICM plus élevé décrit un hôpital dont les cas consomment, en moyenne, davantage de ressources [158,159]. Toutefois, dans un centre universitaire américain, les poids DRG n'expliquaient qu'environ 37 % de la variabilité des coûts directs d'hospitalisation [158].

Dans une méta-analyse de 29 études, le paiement fondé sur les DRG a été associé à une diminution de la durée de séjour, sans effet statistiquement significatif sur les réhospitalisations à 30 jours ni sur la mortalité [160].

4.6.2 Variables et étapes du groupage dans la RO DRG v.1.1

Le manuel RO DRG v.1.1 énumère les variables utilisées pour le groupage : les diagnostics, les actes, le sexe, l'âge, le mode de sortie, la durée de séjour, le poids à l'admission (chez les

nouveau-nés) et l'hospitalisation de jour [139]. Le groupeur, un logiciel dédié, parcourt les étapes dans l'ordre : contrôles démographiques et cliniques, attribution de la catégorie majeure de diagnostic (CMD), traitement pré-CMD, répartition en partitions chirurgicale, "autre" et médicale, attribution du groupe DRG (GDRG), calcul des niveaux de complications et de comorbidités et, enfin, attribution du DRG [139] (figure 4.4).

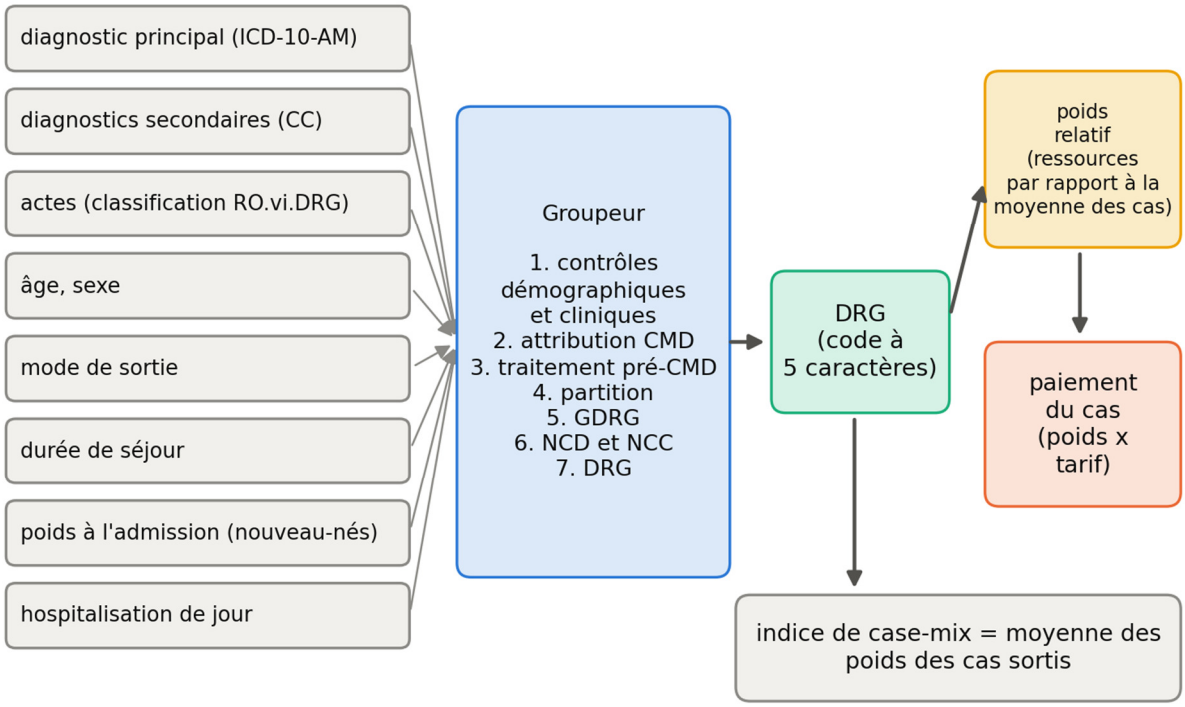


Figure 4.4. Circuit du groupage DRG : des données de sortie au groupe, au poids relatif et au paiement. Schéma simplifié d'après le manuel RO DRG v.1.1 [139] et [158].

Les contrôles cliniques vérifient, entre autres, que le diagnostic principal n'est pas un code de manifestation ou un code de cause externe [139]. Le traitement pré-CMD retire du circuit habituel les cas très coûteux, par exemple la transplantation, l'ECMO, la trachéostomie ou la ventilation mécanique de plus de 95 heures [139]. Les diagnostics secondaires reçoivent un niveau de complication ou de comorbidité de 0 à 4 (de 0 à 3 pour les épisodes médicaux), et ces valeurs servent à calculer le niveau de complexité clinique du patient [139]. Le code DRG comporte cinq caractères : la catégorie majeure, le type (1 chirurgical, 2 autre, 3 médical), le numéro d'ordre et le degré de complexité [139]. Un cas avec un âge invalide ou avec un diagnostic principal incompatible avec le sexe aboutit dans l'un des groupes d'erreur, par exemple 93010 "non groupable" [139].

La conséquence pour le médecin est directe. Un diagnostic secondaire omis peut faire descendre le cas dans un groupe de poids inférieur, et un diagnostic principal mal choisi peut déplacer le cas dans une autre catégorie majeure.

4.6.3 Risques : sous-codage, surcodage, qualité des données

Les données administratives codées ne sont pas le reflet exact de la réalité clinique. Une revue systématique de 280 études de validation a trouvé de larges intervalles interquartiles pour l'exactitude des codes CIM : 44,0-91,4 % pour la sensibilité et 59,4-91,1 % pour la valeur prédictive positive [161]. Dans un hôpital universitaire américain, des algorithmes appliqués aux

données structurées du dossier de santé électronique ont repéré la maladie sans le code correspondant dans 32,9 % des hospitalisations, et 5,8 % des hospitalisations auraient pu recevoir un groupe DRG supérieur [162].

À l'inverse, le **surcodage** (*upcoding*) consiste à coder un cas à un niveau de sévérité supérieur au niveau réel, pour obtenir un paiement plus élevé. Il peut s'agir d'une fraude, mais il peut aussi refléter une documentation plus complète de la sévérité réelle [163]. Dans cinq États américains, les sorties relevant du groupe MS-DRG à l'intensité de codage la plus élevée ont augmenté de 41 % entre 2011 et 2019. Après ajustement sur les caractéristiques des patients, les auteurs estiment que, sans changement du comportement de codage, la hausse aurait été de 13 % [163].

4.7 Le codage assisté par ordinateur et l'intelligence artificielle

Le codage manuel est lent et coûteux. Une revue systématique de 54 études publiées entre 2019 et 2024 décrit le passage de l'apprentissage automatique classique à l'apprentissage profond, avec des réseaux convolutifs, récurrents, de type transformer et hybrides. Les modèles obtiennent de meilleurs résultats sur les codes fréquents que sur l'ensemble des codes, et les codes rares restent un problème non résolu [164].

Les grands modèles de langage (*large language models*, LLM) semblent à première vue bien adaptés, puisqu'ils "lisent" le texte clinique. Les évaluations montrent autre chose. Dans un test portant sur plus de 27 000 codes uniques de diagnostic et d'acte extraits du dossier électronique du système de santé Mount Sinai, les modèles recevaient la description officielle de chaque code et devaient générer le code. Le meilleur modèle, GPT-4, a reproduit exactement 45,9 % des codes ICD-9-CM, 33,9 % des codes ICD-10-CM et 49,8 % des codes CPT [165]. Sur 300 résumés de sortie, même les modèles de raisonnement ont prédit le bon code CIM-9 dans 45,3 % des cas au plus, et le codage a été la tâche la moins bien réussie parmi celles qui ont été testées [166]. Dans un test sur des vignettes cliniques synthétiques, les codes SNOMED CT proposés par les LLM étaient systématiquement invalides, alors que les libellés textuels semblaient plausibles [167]. On voit ici un risque typique : le modèle génère un code qui a l'air correct, mais qui n'existe pas.

L'aide utile vient de systèmes qui proposent des codes, la décision revenant à l'humain. Dans une étude randomisée croisée menée en Suède et en Norvège, l'outil Easy-ICD a réduit de 46 % le temps médian de codage des notes cliniques longues. L'exactitude a augmenté (de 62 % à 67 % pour les notes longues), mais la différence n'était pas statistiquement significative [168]. L'étude n'a porté que sur 15 participants, dans un cadre en ligne, et les auteurs demandent des évaluations dans le flux de travail réel de l'hôpital [168].

Pièges fréquents

- Coder directement de mémoire ou à partir d'une liste de "codes usuels", sans vérification dans la liste tabulaire. Les notes d'exclusion changent souvent le code.
- Confondre la cause terminale et la cause initiale du décès. La pneumonie inscrite sur la première ligne du certificat n'est pas automatiquement la cause qui est comptée.
- Utiliser un code d'extension CIM-11 seul ou en première position du cluster. Les extensions s'attachent toujours à un code souche.
- Croire que les tables de correspondance CIM-10-CIM-11 convertissent automatiquement les données. L'OMS les recommande pour la comparaison, non pour la conversion.
- Se fier à un code proposé par un LLM sans vérification dans le navigateur officiel. Le code peut être plausible et pourtant inexistant.
- Omettre les diagnostics secondaires à la sortie, ce qui influe sur le groupe DRG et sur les statistiques de l'hôpital.

À retenir

- Une classification regroupe les cas en classes exhaustives et mutuellement exclusives. Une nomenclature nomme les concepts aussi précisément que possible.
- CIM-10 : chapitre, bloc, catégorie à 3 caractères, sous-catégorie à 4 caractères. On cherche dans l'index, on vérifie dans la liste tabulaire.
- La cause initiale du décès est la maladie qui a déclenché la chaîne d'événements ou les circonstances de l'accident ou de la violence.
- La CIM-11 est en vigueur depuis le 1er janvier 2022. La MMS est la linéarisation statistique extraite de la Fondation.
- Dans la CIM-11, "&" ajoute des codes d'extension, "/" relie deux codes souches.
- La Roumanie code les diagnostics des patients hospitalisés en ICD-10-AM, 3e éd., et groupe les cas selon la RO DRG v.1.1 (669 groupes).
- SNOMED CT, LOINC, l'ATC et MeSH ont des finalités différentes : concepts cliniques, analyses, médicaments, littérature.
- Les modèles de langage généralistes se trompent fréquemment de code. Tout code proposé automatiquement doit être vérifié.

Questions d'autoévaluation

1. (QCS) Dans la CIM-10, le code I21.0 représente :

a) un chapitre
b) un bloc de catégories
c) une catégorie à trois caractères
d) une sous-catégorie à quatre caractères
e) un code d'extension

2. (QCS) Un certificat de décès indique : (a) pneumonie, due à (b) une fracture de la hanche, due à (c) un trébuchement sur un tapis. Selon le guide de l'OMS, le point de départ de la séquence est :

a) la pneumonie
b) la fracture de la hanche
c) le trébuchement sur le tapis (l'événement externe)
d) l'âge du patient
e) l'insuffisance respiratoire terminale

3. (QCS) Dans le cluster CIM-11 BA41.0&XA7RE3&XA7NQ7, les codes XA7RE3 et XA7NQ7 sont :

a) des codes souches désignant d'autres maladies associées
b) des codes d'extension qui ajoutent des détails anatomiques
c) des codes CIM-10 conservés par souci de compatibilité
d) des codes d'interventions ICHI
e) des codes SNOMED CT

4. (QCM) Concernant le système DRG en Roumanie, au 27 septembre 2026, les affirmations correctes sont :

a) les diagnostics sont codés en ICD-10-AM, 3e édition
b) la classification en vigueur est la RO DRG v.1.1, qui compte 669 groupes
c) les groupes sont déterminés uniquement d'après le diagnostic principal, sans autres variables
d) l'âge, le sexe et le mode de sortie sont des variables de groupage
e) la CIM-11 a officiellement remplacé l'ICD-10-AM dans les hôpitaux depuis 2022

5. (QCM) Concernant les terminologies médicales, les affirmations correctes sont :

a) le nom complet d'un terme LOINC comporte cinq ou six parties, dont le composant, le système et l'échelle
b) dans la classification ATC, la metformine a le code A10BA02, au cinquième niveau
c) MeSH est utilisé principalement pour la facturation des hôpitaux
d) SNOMED CT permet des expressions postcoordonnées
e) l'UMLS est une classification statistique des causes de décès

CHAPITRE 5

Chapitre 5. Éthique, protection des données et sécurité de l'information en santé

Objectifs

À la fin de ce chapitre, l'étudiant sera capable :

- d'appliquer les quatre principes de l'éthique biomédicale à des situations concrètes d'utilisation des données de santé
- de reconnaître les données concernant la santé au sens du RGPD et d'identifier la base juridique d'un traitement
- de distinguer l'anonymisation de la pseudonymisation et d'estimer le risque de réidentification d'un jeu de données
- de décrire la triade confidentialité-intégrité-disponibilité et les mesures de base qui la protègent dans un hôpital
- d'expliquer comment apparaît le biais algorithmique et comment il peut affecter l'équité des soins
- de citer correctement, d'éviter le plagiat et de déclarer l'utilisation d'outils d'intelligence artificielle (IA) générative

Les données d'un patient circulent aujourd'hui entre l'hôpital, le laboratoire, le médecin de famille, la caisse d'assurance maladie et, de plus en plus souvent, des bases de données de recherche. Que doit le médecin au patient lorsque ces données deviennent électroniques ? La réponse comporte un volet éthique, un volet juridique, construit autour du règlement général sur la protection des données (RGPD), et un volet technique : sans sécurité de l'information, les obligations éthiques et légales restent des déclarations d'intention.

5.1 Principes éthiques appliqués aux données de santé

5.1.1 Les quatre principes

Les valeurs couramment admises en éthique médicale sont résumées par les quatre principes de Beauchamp et Childress [169]. Ils se transposent directement dans l'univers des données.

L'**autonomie** signifie que la personne décide de ce qu'il advient des informations qui la concernent : elle sait qui voit ses données et dans quel but, et son accord, son refus ou son retrait sont respectés. La **bienfaisance** exige que les données soient utilisées dans l'intérêt du patient et de la collectivité : un dossier électronique complet peut réduire les erreurs, et les données agrégées permettent des études qui améliorent les traitements. La **non-malfaisance** (*primum non nocere*) oblige à prévenir le tort causé par la divulgation, par la discrimination ou par des décisions prises sur des données erronées. La **justice** concerne la répartition équitable des bénéfices et des risques, par exemple le fait qu'un algorithme entraîné sur une seule population ne désavantage pas d'autres groupes (section 5.5).

Les principes entrent souvent en conflit. Un exemple récent est celui des données médicales synthétiques, générées par l'IA pour se substituer aux données réelles. Friedman et Smajdor montrent que, lorsque la protection de la vie privée (*privacy*) est traitée comme une valeur suprême, elle peut entrer en conflit avec l'autonomie et avec la bienfaisance, de sorte que l'avantage éthique des données synthétiques ne peut être présumé sans une analyse attentive [169].

5.1.2 Confidentialité et secret professionnel

Les termes "vie privée" et "confidentialité" ne sont pas synonymes. La vie privée se rapporte à la personne, la confidentialité à l'information : celui qui suit un patient jusqu'à l'entrée d'une clinique porte atteinte à sa vie privée, celui qui ouvre son dossier sans en avoir le droit viole la confidentialité [170]. Le même chapitre du manuel de Shortliffe et Cimino formule l'idée de base à l'intention de l'étudiant en médecine : protéger la confidentialité n'est pas une faveur faite au patient, mais un devoir, quels que soient la maladie et le support sur lequel l'information est stockée [170]. La raison est aussi pratique : celui qui craint la divulgation cherche moins souvent une aide médicale, ce qui augmente le risque que les infections se propagent et que les maladies restent non traitées [170].

En Roumanie, la loi n° 46/2003 sur les droits du patient (Legea drepturilor pacientului) prévoit que toutes les informations relatives à l'état du patient, aux résultats des examens, au diagnostic, au pronostic et au traitement sont confidentielles, même après le décès (art. 21). Elles ne peuvent être communiquées qu'avec le consentement explicite du patient ou si la loi l'exige expressément (art. 22), mais le consentement n'est plus requis lorsque les informations sont nécessaires à d'autres prestataires accrédités qui participent au traitement (art. 23) [171]. Le RGPD relie explicitement le traitement des données médicales au secret professionnel : les données peuvent être traitées aux fins du diagnostic et des soins par un professionnel soumis à une obligation de secret professionnel ou sous sa responsabilité (art. 9, par. 3) [118]. La confidentialité n'est pas absolue. L'équipe soignante a un accès légitime au dossier, et certaines maladies transmissibles sont déclarées aux autorités de santé publique [170].

5.1.3 Le consentement éclairé, du formulaire à la plateforme

Le consentement éclairé suppose une information adéquate, une compréhension et une décision libre. Pour l'intervention médicale, la loi n° 46/2003 lui consacre un chapitre (chap. III) [171]. En recherche, la discussion s'est déplacée vers la forme électronique (*eConsent*). Une revue systématique de 35 études totalisant 13 281 participants n'a trouvé aucune étude dans laquelle le formulaire papier donnait de meilleurs résultats que la version électronique en matière de compréhension, d'acceptabilité ou d'utilisabilité [172].

La réutilisation des données cliniques pour la recherche a donné naissance à deux modèles alternatifs. Le **consentement large** (*broad consent*) couvre une catégorie de recherches futures, et non une seule étude. En Allemagne, l'Initiative d'informatique médicale a élaboré un consentement large national conforme au RGPD, assorti de garanties obligatoires : évaluation indépendante de chaque projet, protection contre la divulgation involontaire et transparence publique [173]. Le **consentement dynamique** s'appuie sur une plateforme numérique grâce à laquelle le participant met à jour ses préférences au fil du temps et peut retirer son accord. Une revue de portée (*scoping review*) a identifié comme facteurs favorables les options multiples et la communication continue, et comme obstacles la lassitude liée au consentement, la fracture numérique, la complexité technique et les risques de sécurité [174]. Une analyse juridique de 2026 montre que l'espace européen des données de santé déplace la gouvernance de l'utilisation secondaire vers une autorisation

institutionnelle, délivrée par des organismes d'accès aux données, ce qui réduit le rôle du consentement individuel pour de nombreuses utilisations à grande échelle [175].

5.2 Le RGPD pour le médecin

5.2.1 Que sont les données concernant la santé

Le règlement (UE) 2016/679 (RGPD) définit les **données à caractère personnel** comme toute information se rapportant à une personne physique identifiée ou identifiable, directement ou indirectement, par exemple par référence à un nom, à un numéro d'identification, à des données de localisation ou à des éléments spécifiques propres à son identité physique, physiologique ou génétique (art. 4, point 1). Les **données concernant la santé** sont les données relatives à la santé physique ou mentale d'une personne, y compris la prestation de services de soins de santé, qui révèlent des informations sur son état de santé (art. 4, point 15) [118]. La définition est large : un rendez-vous en oncologie ou une ordonnance révèlent l'état de santé.

Les données concernant la santé font partie des catégories particulières, dont le traitement est en principe interdit (art. 9, par. 1). L'interdiction n'est levée que dans les situations énumérées à l'art. 9, par. 2. Pour le médecin comptent avant tout le point h (médecine préventive ou médecine du travail, diagnostics médicaux, prise en charge sanitaire et gestion des services de soins de santé), le point i (motifs d'intérêt public dans le domaine de la santé publique), le point j (recherche scientifique, avec les garanties de l'art. 89), le point c (intérêts vitaux d'une personne qui se trouve dans l'incapacité de donner son consentement) et le point a (consentement explicite) [118]. En outre, tout traitement a besoin d'une base juridique au titre de l'art. 6, par. 1, par exemple une obligation légale du responsable du traitement ou une mission d'intérêt public. La figure 5.1 résume ce raisonnement.

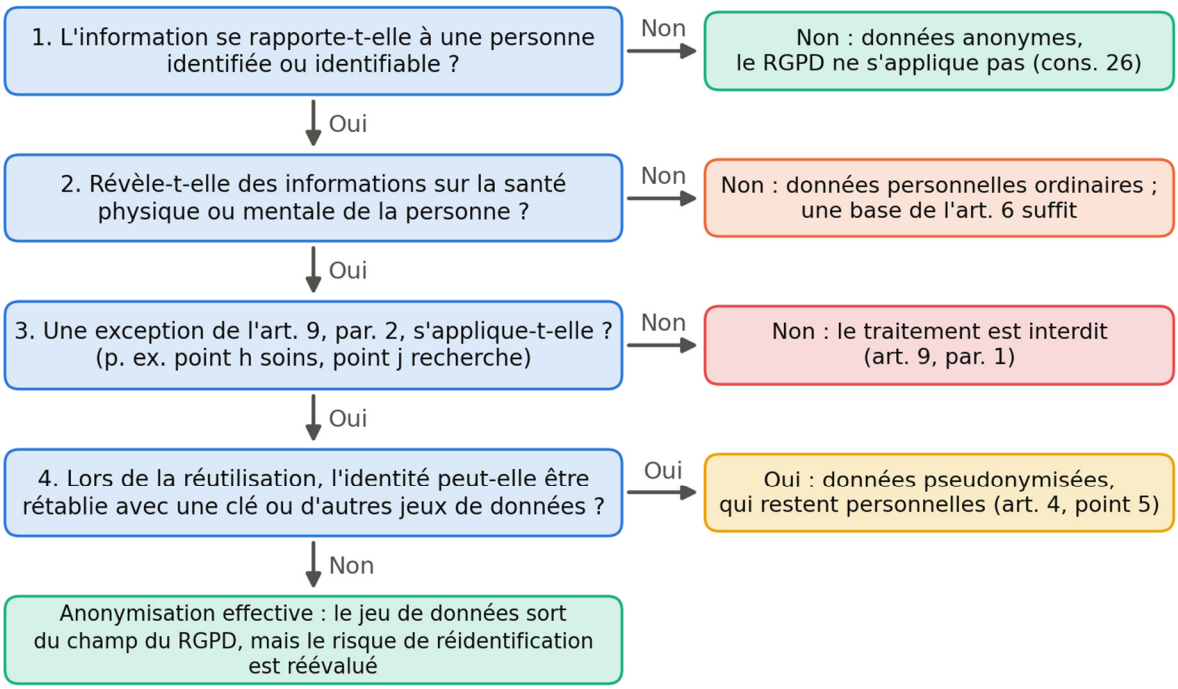


Figure 5.1. Arbre de décision pour qualifier une information : données personnelles, données concernant la santé, base juridique, données pseudonymisées ou anonymes. Élaboré d'après le règlement (UE) 2016/679, art. 4, art. 9 et considérant 26 [118].

Un détail souvent ignoré : le consentement n'est pas la base habituelle des soins courants. L'hôpital traite les données du patient hospitalisé parce qu'elles sont nécessaires au diagnostic et au traitement (art. 9, par. 2, point h), et non parce que le patient a coché une case. Le consentement à l'intervention médicale, exigé par la loi n° 46/2003, est une institution distincte.

5.2.2 Principes et droits

L'art. 5 énumère les principes relatifs au traitement : licéité, loyauté et transparence, limitation des finalités, minimisation des données, exactitude, limitation de la conservation, intégrité et confidentialité, auxquels s'ajoute la responsabilité du responsable du traitement, tenu de démontrer la conformité [118]. La **limitation des finalités** signifie que les données collectées pour les soins ne sont pas utilisées ultérieurement d'une manière incompatible, la recherche assortie de garanties appropriées étant considérée comme compatible. La **minimisation des données** (*data minimisation*) exige que seules les variables nécessaires figurent dans un export destiné à une étude. Les droits du patient en tant que personne concernée sont présentés dans le tableau 5.1.

Tableau 5.1. Droits de la personne concernée selon le RGPD, avec des exemples tirés de la pratique médicale. Source : [118,176].

Droit (article du RGPD)	Ce que le patient peut demander	Remarque pour le médecin
Information (art. 13-14)	qui traite les données, dans quel but, pendant combien de temps	la notice d'information à l'admission
Accès (art. 15)	une copie des données du dossier	aussi loi n° 46/2003, art. 24
Rectification (art. 16)	la correction des données inexactes	les erreurs cliniques se corrigent en laissant une piste d'audit
Effacement (art. 17)	l'effacement des données	limité par l'obligation légale d'archivage
Limitation (art. 18)	le blocage temporaire du traitement	p. ex. pendant la durée d'une contestation
Portabilité (art. 20)	les données dans un format structuré, lisible par machine	s'applique aux traitements fondés sur le consentement ou sur un contrat
Opposition (art. 21)	l'arrêt d'un traitement	dérogations pour la recherche (loi n° 190/2018, art. 8)
Décisions automatisées (art. 22)	ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé	loi n° 190/2018, art. 3, al. 1

5.2.3 DPO, analyse d'impact et violations de données

Un hôpital public doit désigner un **délégué à la protection des données** (*data protection officer*, DPO), à la fois parce qu'il est un organisme public et parce qu'il traite à grande échelle des catégories particulières de données (art. 37, par. 1, points a et c) [118]. Le DPO informe et conseille le responsable du traitement et les employés, contrôle le respect du règlement et fait office de point de contact avec l'autorité de contrôle, qui est en Roumanie l'Autorité nationale de surveillance du traitement des données à caractère personnel (Autoritatea Națională de Supraveghere a Prelucrării Datelor cu Caracter Personal, ANSPDCP).

Avant un traitement susceptible d'engendrer un risque élevé, le responsable du traitement effectue une **analyse d'impact relative à la protection des données** (AIPD, en anglais *data protection impact assessment*, DPIA). Le règlement l'exige explicitement pour le traitement à grande échelle de catégories particulières de données (art. 35, par. 3, point b) [118], par exemple lors de la création d'un registre régional de maladie.

Lorsque survient une **violation de données à caractère personnel** (accès non autorisé, perte, chiffrement par un rançongiciel ou *ransomware*), le responsable du traitement la notifie à l'autorité dans les meilleurs délais et, si possible, 72 heures au plus tard après en avoir pris connaissance. Il n'est dispensé de notification que si la violation n'est pas susceptible d'engendrer un risque pour les personnes (art. 33) [118]. Si le risque est élevé, les patients concernés sont également informés (art. 34). Pour le médecin, la règle pratique est simple : un ordinateur portable perdu, un courriel contenant des résultats envoyé au mauvais destinataire ou un accès suspect se signalent immédiatement au DPO, car le délai court à partir du moment où l'organisation en a connaissance.

5.2.4 La loi n° 190/2018

La loi n° 190/2018 (Legea nr. 190/2018) établit les mesures nationales de mise en œuvre du RGPD. L'art. 3, al. 1, n'autorise le traitement des données génétiques, biométriques ou concernant la santé en vue d'une prise de décision automatisée ou d'un profilage qu'avec le consentement explicite de la personne ou en vertu de dispositions légales expresses, assorties de mesures de protection. L'art. 3, al. 2, interdit que les données traitées aux fins de la protection de la santé publique soient utilisées ultérieurement, à d'autres fins, par des tiers. L'art. 4 encadre le traitement du code numérique personnel (CNP). L'art. 8 permet, pour la recherche scientifique, des dérogations aux droits d'accès, de rectification, de limitation et d'opposition, uniquement si leur exercice risque de rendre la recherche impossible ou de l'entraver gravement, et uniquement avec les garanties de l'art. 89 du RGPD. Pour les organismes publics, donc aussi pour les hôpitaux publics, les art. 13-14 prévoient d'abord un avertissement assorti d'un plan de remédiation, puis, si ce plan n'est pas exécuté, des amendes de 10 000 à 100 000 lei, et de 10 000 à 200 000 lei en cas de violation des principes de base du traitement et des droits des personnes concernées [176].

5.2.5 Anonymisation et pseudonymisation

La **pseudonymisation** est le traitement de données effectué de telle façon que celles-ci ne puissent plus être attribuées à une personne sans avoir recours à des informations supplémentaires, conservées séparément et protégées par des mesures techniques et organisationnelles (art. 4, point 5). Les données pseudonymisées restent des données à caractère personnel. En revanche, les données **anonymes** ne concernent plus une personne identifiable et échappent au RGPD (considérant 26), mais le critère est exigeant : il faut prendre en considération l'ensemble des moyens raisonnablement susceptibles d'être utilisés par le responsable du traitement ou par toute autre personne, y compris le coût, le temps nécessaire et la technologie disponible [118]. Les différences sont résumées dans le tableau 5.2.

Tableau 5.2. Anonymisation et pseudonymisation, comparaison pratique. Source : [118,177].

Critère	Pseudonymisation	Anonymisation
Définition	identité remplacée par un code, clé conservée séparément	lien avec la personne supprimé de manière irréversible
Statut au regard du RGPD	données personnelles (art. 4, point 5)	hors du champ du RGPD (cons. 26)
Réversibilité	oui, par le détenteur de la clé	non (si elle est réussie)
Utilité pour la recherche	élevée, permet le chaînage et le suivi dans le temps	réduite par la généralisation et la suppression
Exemple	registre avec un code d'étude à la place du CNP	tableau publié avec des âges par tranches, sans données rares

La pseudonymisation est recommandée par le RGPD comme mesure de sécurité (art. 32) et comme garantie pour la recherche (art. 89) [118]. En pratique, la clé de correspondance est conservée séparément, par une personne ou un service de confiance, et le chercheur ne reçoit que le code.

L'anonymisation réelle est plus difficile à obtenir qu'il n'y paraît. Un diagnostic rare combiné à l'âge, au sexe et au code postal peut fonctionner comme une empreinte du patient, même après la suppression du nom [170]. Une revue des techniques d'anonymisation a montré, au moyen d'une attaque par reconstruction, que l'anonymisation est nécessaire, mais non suffisante pour protéger les patients [177]. Les images médicales comportent des risques qui leur sont propres. Sur le jeu de données public ChestX-ray14 (112 120 radiographies provenant de 30 805 patients), un réseau de neurones a reconnu si deux radiographies thoraciques appartenaient à la même personne avec une aire sous la courbe (AUC) de 0,994, même lorsque plus de dix ans séparaient les examens [178].

Il ne faut pas pour autant surestimer le risque. Sur les données de plus d'un million de patients, un modèle qui tenait compte des ressources réelles de l'attaquant a montré que, par rapport au scénario le plus pessimiste (risque médian de réidentification de 0,987), la réduction médiane du risque était de 90,1 % lorsque l'attaquant utilisait des listes électorales et de 100 % lorsqu'il utilisait des publications sur les réseaux sociaux [179].

5.3 L'espace européen des données de santé

Le règlement (UE) 2025/327 relatif à l'espace européen des données de santé (*European Health Data Space*, EHDS) est entré en vigueur le 26 mars 2025 et s'applique, de manière générale, à partir du 26 mars 2027 [23]. Il complète le RGPD et distingue deux utilisations des mêmes données.

L'**utilisation primaire** désigne l'utilisation des données pour les soins, y compris dans un autre État membre. Le patient a un droit d'accès aux catégories prioritaires de données (art. 3) et peut restreindre l'accès des professionnels à certaines informations (art. 8), et les États peuvent instaurer un droit d'opposition à l'utilisation primaire (art. 10). Les catégories prioritaires sont le résumé du patient, les prescriptions et dispensations électroniques, les examens d'imagerie avec leurs comptes rendus, les résultats d'examens et les rapports de sortie (art. 14). Les obligations s'appliquent à partir du 26 mars 2029 pour le résumé du patient et les prescriptions et dispensations électroniques, et à partir du 26 mars 2031 pour les autres catégories [23]. Les systèmes de dossier électronique doivent intégrer un composant européen d'interopérabilité et un composant européen de journalisation (art. 25), c'est-à-dire précisément le journal d'audit dont il est question dans la section 5.4 [23]. L'infrastructure MyHealth@EU et les échanges transfrontaliers sont décrits au chapitre 8.

L'**utilisation secondaire** désigne la réutilisation des données pour la recherche, la santé publique, la statistique, l'élaboration des politiques ou l'enseignement (art. 53). Le chapitre IV du règlement s'applique à partir du 26 mars 2029, et l'art. 51 énumère 17 catégories minimales de données, des données des dossiers électroniques jusqu'aux registres et aux biobanques [23]. Chaque État désigne un ou plusieurs **organismes responsables de l'accès aux données de santé** (*health data access bodies*), qui délivrent des autorisations (art. 55, art. 68), ainsi qu'un point de contact national connecté à la plateforme européenne HealthData@EU (art. 75). Les données ne sont mises à disposition que dans la mesure nécessaire à la finalité, sous forme anonymisée lorsque la finalité peut être atteinte de cette manière (art. 66), et uniquement dans un

environnement de traitement sécurisé, qui limite l'accès aux personnes autorisées et réduit au minimum le risque de copie ou d'extraction non autorisée (art. 73) [23].

Deux garanties sont importantes pour le patient. L'art. 54 interdit d'utiliser les données pour prendre des décisions préjudiciables à la personne, pour l'exclure d'une assurance ou d'un crédit, à des fins publicitaires ou pour développer des produits nocifs, tels que les produits du tabac. L'art. 71 donne à toute personne le droit de refuser, à tout moment et sans justification, l'utilisation secondaire de ses données de santé à caractère personnel, et cette décision est réversible [23].

5.4 La sécurité de l'information

5.4.1 La triade confidentialité-intégrité-disponibilité

La sécurité de l'information se décrit par trois propriétés. La **confidentialité** signifie que les données ne sont accessibles qu'aux personnes autorisées. L'**intégrité** signifie que les données sont exactes et complètes et qu'elles n'ont pas été modifiées sans autorisation : une dose modifiée dans une prescription est un problème d'intégrité. La **disponibilité** signifie que les données et les systèmes peuvent être utilisés lorsque c'est nécessaire, et en santé, "lorsque" peut vouloir dire à trois heures du matin, aux urgences. Le RGPD reprend la même triade et exige du responsable du traitement des mesures garantissant la confidentialité, l'intégrité, la disponibilité et la résilience constantes des systèmes (art. 32) [118]. Il existe des compromis entre les trois propriétés. Toute restriction d'accès a un coût, parfois explicite (équipements et logiciels de sécurité), parfois implicite, sous la forme du temps demandé aux utilisateurs [170].

5.4.2 Menaces

Le premier rapport de l'Agence de l'Union européenne pour la cybersécurité (ENISA) consacré à la santé a analysé 215 incidents signalés publiquement entre janvier 2021 et mars 2023. Les prestataires de soins de santé ont été la cible dans 53 % des incidents, et les hôpitaux dans 42 %. Les rançongiciels, qui chiffrent les données et exigent une rançon, ont représenté 54 % des menaces, et les données des patients, y compris les dossiers électroniques, ont été l'actif le plus visé (30 %) (figure 5.2) [180]. Le même rapport note que 27 % seulement des organisations de santé interrogées disposaient d'un programme dédié de défense contre les rançongiciels et que 40 % n'avaient aucun programme de sensibilisation pour le personnel extérieur au service informatique [180].

La porte d'entrée est souvent un message. Dans les hôpitaux étudiés par Jalali et collaborateurs (397 employés dont les réponses à un questionnaire ont été comparées aux clics réels lors de campagnes de test), l'intention déclarée de respecter les règles ne prédisait pas le comportement, et le seul facteur associé au clic sur un lien d'hameçonnage (*phishing*) était la charge de travail [181]. La menace interne (*insider*) est tout aussi réelle. Les systèmes qui facilitent l'accès des cliniciens le facilitent aussi pour l'employé curieux qui n'a pas besoin de l'information [170]. Dans une enquête menée auprès de 41 professionnels de centres médicaux espagnols, 97,6 % s'authentifiaient uniquement par mot de passe, et plus de 40 % envoyaient des données sensibles par courriel personnel ou par des applications de messagerie [182].

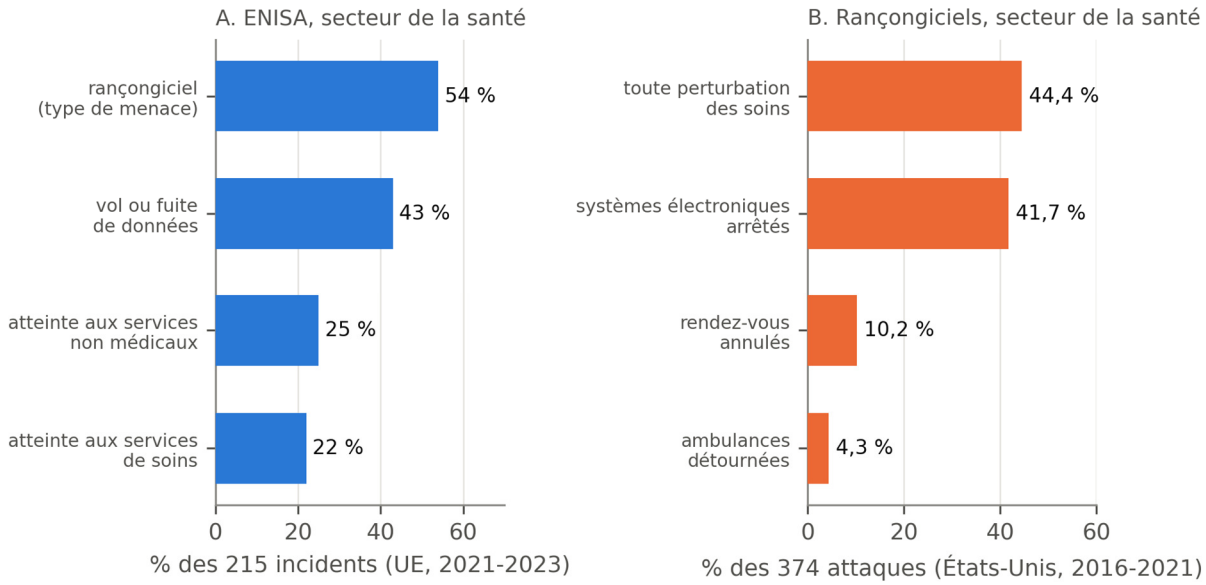


Figure 5.2. Incidents de sécurité en santé, données réelles. A : part des rançongiciels et des principales conséquences dans 215 incidents survenus dans l'UE et dans les pays voisins, janvier 2021-mars 2023 (les pourcentages ne s'additionnent pas, un incident pouvant relever de plusieurs catégories). B : conséquences sur les soins de 374 attaques par rançongiciel contre des établissements de santé aux États-Unis, 2016-2021. Données issues de [180,183].

5.4.3 Quand l'attaque devient un événement clinique

Entre 2016 et 2021, 374 attaques par rançongiciel contre des établissements de santé aux États-Unis ont exposé les données de près de 42 millions de patients, et leur nombre annuel a doublé, passant de 43 à 91. Près de la moitié d'entre elles (44,4 %) ont perturbé les soins, le plus souvent par l'arrêt des systèmes électroniques (41,7 %), l'annulation de rendez-vous (10,2 %) ou le détournement des ambulances (4,3 %). Au fil du temps, les attaques ont été de moins en moins souvent résolues par une restauration à partir de sauvegardes [183].

Les effets se propagent aussi aux hôpitaux voisins. Lorsqu'un réseau hospitalier de la région de San Diego a été attaqué en mai 2021, les services des urgences d'un centre universitaire voisin, non touché par l'attaque, ont accueilli en moyenne 251,4 patients par jour, contre 218,4 auparavant, et les déclenchements de l'alerte accident vasculaire cérébral sont passés de 59 à 102 [184]. Dans les deux hôpitaux de ce même centre, les arrêts cardiorespiratoires sont passés de 21 à 38 pendant la période de l'attaque, et la survie avec un état neurologique favorable après un arrêt cardiaque extrahospitalier a chuté de 40,0 % à 4,5 % [185]. Il s'agit d'études observationnelles, portant sur un petit nombre d'événements, qui montrent des associations et non une causalité. Elles aboutissent néanmoins à une conclusion commune : une cyberattaque doit être traitée comme une catastrophe régionale.

La situation en Roumanie en 2026

- En février 2024, une attaque par le rançongiciel Backmydata, de la famille Phobos, a chiffré les serveurs sur lesquels fonctionnait le système d'information hospitalier HIPOCRATE dans plusieurs hôpitaux. Le 12 février, la Direction nationale de la cybersécurité (Directoratul Național de Securitate Cibernetică, DNSC) confirmait 21 hôpitaux touchés, dont l'Institut des maladies cardiovasculaires de Timișoara, et 79 autres établissements avaient été déconnectés d'Internet pour investigation [186].
- Les attaquants ont exigé 3,5 BTC (environ 157 000 euros). La DNSC a recommandé de ne pas payer la rançon et de ne pas entrer en contact avec les attaquants. Le 15 février, l'alerte de la DNSC faisait état de 26 hôpitaux [186].
- La plupart des hôpitaux touchés disposaient de sauvegardes récentes, datant de 1 à 3 jours, à une seule exception près (12 jours), ce qui, selon la DNSC, pouvait faciliter la restauration des données et des services [186].
- Le préambule de l'ordonnance d'urgence du Gouvernement (OUG) n° 155/2024, qui transpose NIS2, invoque cet incident, qui "a touché 26 hôpitaux, au niveau national, par l'intermédiaire d'un fournisseur de services gérés" (notre traduction) [187].

5.4.4 NIS2 et sa transposition en Roumanie

La directive (UE) 2022/2555 (NIS2) s'applique depuis le 18 octobre 2024 et range la santé parmi les secteurs hautement critiques : prestataires de soins de santé, laboratoires de référence, fabricants de médicaments et de dispositifs médicaux critiques [188]. Les grandes entités de ces secteurs sont des entités "essentielles" [188]. Elles doivent appliquer des mesures de gestion des risques (art. 21), et les organes de direction approuvent ces mesures et peuvent être tenus responsables des manquements (art. 20) [180,188]. Un incident important se signale en trois temps : alerte précoce dans les 24 heures, notification dans les 72 heures et rapport final dans un délai d'un mois au plus (art. 23) [188]. En Roumanie, la directive a été transposée par l'OUG n° 155/2024, publiée le 31 décembre 2024. La DNSC est l'autorité compétente et l'équipe nationale de réponse aux incidents (CSIRT), et les délais de 24 heures, de 72 heures et d'un mois sont repris à l'art. 15, al. 7 [187]. Une même violation peut donc déclencher deux notifications parallèles : l'une à la DNSC, au titre de NIS2, l'autre à l'ANSPDCP, au titre du RGPD, si des données personnelles sont touchées.

5.4.5 Mesures de base

La norme ISO 27799, dont la troisième édition a été publiée en décembre 2025, adapte à la santé les mesures de sécurité de l'ISO/IEC 27002:2022 [189]. Les recommandations ci-dessous, résumées dans la figure 5.3, sont tirées du projet soumis au vote en 2025 [190].

- **Authentification multifacteur.** Pour les systèmes qui traitent des données personnelles de santé, on recommande au moins deux facteurs, par exemple le mot de passe plus un code reçu sur le téléphone ou une carte [190].
- **Moindre privilège.** L'accès est accordé par rôle, limité aux besoins du rôle, et l'utilisateur n'accède au dossier que s'il fait partie de l'équipe soignante ou s'il a une autre mission légitime. L'accès temporaire des étudiants, des internes et des médecins remplaçants est clôturé à leur départ [190].
- **Journaux d'audit.** Chaque consultation et chaque modification sont enregistrées. Les journaux montrent a posteriori qui a accédé au dossier et découragent la curiosité, surtout lorsque les sanctions en cas de manquement sont connues et appliquées [170].

- **Sauvegardes.** Les données de santé sont sauvegardées sous forme chiffrée, et, contre les rançongiciels, on recommande des copies hors ligne ou immuables [190]. Une sauvegarde qui n'a jamais été restaurée à titre d'essai reste une hypothèse.
- **Plan de continuité.** Les formulaires papier et les procédures de travail sans système se préparent à l'avance, et les données consignées pendant l'interruption sont ensuite ressaisies avec des vérifications supplémentaires [190].

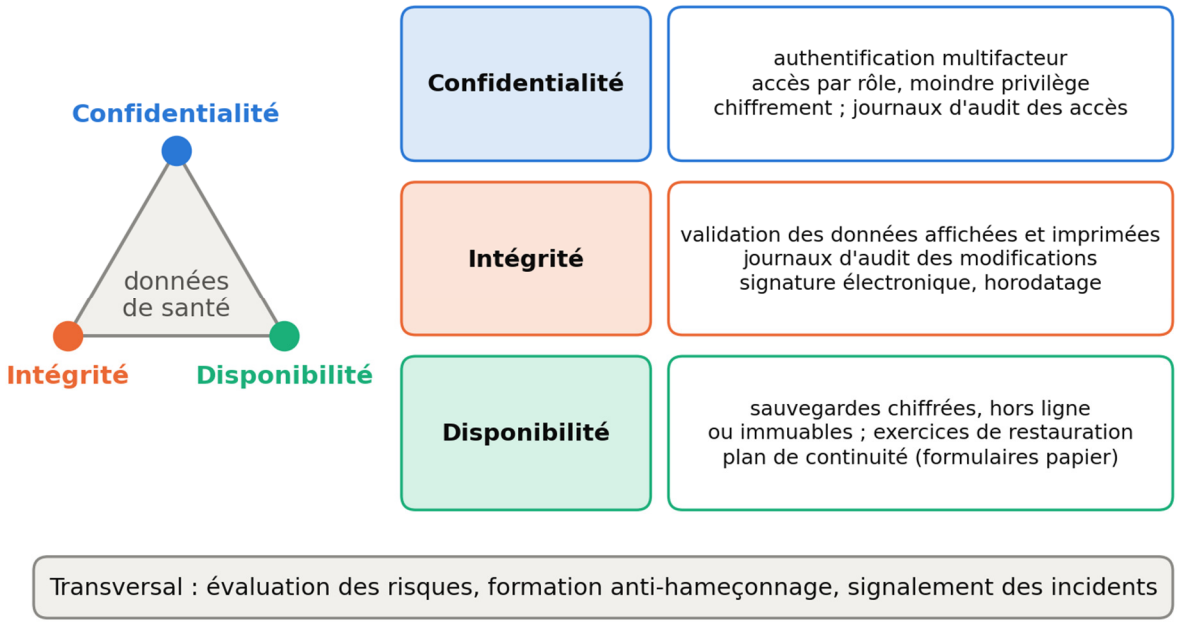


Figure 5.3. La triade confidentialité-intégrité-disponibilité et les mesures de sécurité qui la soutiennent dans un établissement de santé. Élaboré d'après [118,190].

5.5 Équité et biais algorithmique

Un algorithme peut être précis en moyenne et néanmoins inéquitable. Le **biais algorithmique** désigne une erreur systématique dont les conséquences pèsent de manière disproportionnée sur un groupe particulier de patients. Le risque apparaît à chaque étape, du choix du problème et des données d'entraînement jusqu'à l'utilisation après le déploiement, et les modèles peuvent amplifier les inégalités existantes [191].

L'oxymétrie de pouls est l'exemple le plus cité, bien qu'elle ne fasse pas intervenir l'IA. Sur 10 789 paires de mesures de l'Université du Michigan, une saturation artérielle inférieure à 88 %, non détectée par un oxymètre de pouls qui affichait 92-96 % (hypoxémie occulte), est survenue dans 11,7 % des mesures chez les patients noirs, contre 3,6 % chez les patients blancs. Dans la cohorte multicentrique, les proportions étaient de 17,0 % et de 6,2 % [192]. Une méta-analyse de 2024 portant sur 15 études et 207 464 patients hospitalisés a confirmé une prévalence plus élevée de l'hypoxémie occulte chez les patients noirs (rapport de prévalence 1,67, IC à 95 % 1,47-1,90) et a recommandé aux cliniciens une vigilance accrue chez les patients à pigmentation cutanée foncée [193]. La leçon pour l'informatique est directe : un capteur étalonné sur une population homogène produit des données systématiquement erronées pour d'autres groupes, et tout modèle construit sur ces données hérite de l'erreur.

Le deuxième exemple vient de l'imagerie. Des classificateurs de pointe pour les radiographies thoraciques, entraînés sur trois grands jeux de données, ont sous-diagnostiqué de manière

sélective des groupes défavorisés (femmes, patients noirs, patients de faible statut socio-économique), avec des taux plus élevés pour les sous-groupes intersectionnels, par exemple les femmes hispaniques [194]. Les formules classiques peuvent elles aussi être corrigées. Les équations de 2021 pour le débit de filtration glomérulaire estimé n'utilisent plus la race, et celles qui reposent sur la créatinine et la cystatine C se sont révélées plus exactes et ont réduit les écarts entre groupes par rapport aux équations à un seul marqueur [195].

Les mesures de réduction du biais partent des données. Les recommandations de consensus STANDING Together, élaborées avec la contribution de plus de 350 représentants de 58 pays, demandent une documentation transparente de la composition et des limites des jeux de données ainsi que l'évaluation des performances par sous-groupes de population [196]. La détection du biais exige parfois justement des données sensibles, comme l'origine ethnique. Le règlement (UE) 2026/1744, qui modifie le règlement sur l'IA (AI Act), a introduit un art. 4 bis qui permet, dans la stricte mesure nécessaire et avec des garanties, le traitement de catégories particulières de données pour détecter et corriger les biais [197]. Les autres exigences applicables aux systèmes d'IA à haut risque sont traitées au chapitre 13 [22].

5.6 Intégrité académique

5.6.1 Plagiat et autoplagiat

La loi n° 199/2023 de l'enseignement supérieur (Legea învățământului superior) définit le **plagiat** comme la présentation, en tant que création personnelle, de textes, d'idées, de démonstrations, de données, de théories, de résultats ou de méthodes empruntés aux œuvres d'autres auteurs, sans le mentionner et sans renvoi aux sources originales. L'**autoplagiat** est la republication de parties substantielles de ses propres publications antérieures, y compris sous forme de traductions, sans citation appropriée de l'original (art. 169, lettres d et e) [198]. La définition couvre aussi les idées, et pas seulement les phrases copiées.

Le phénomène n'est pas rare, même dans la formation des professionnels de santé. Une méta-analyse de 31 études portant sur 9 175 étudiants en soins infirmiers a établi que le plagiat était la forme la plus fréquente de fraude académique (55,3 %). La forme la plus répandue était la paraphrase sans renvoi à la source (39,5 %), et le plagiat était un facteur prédictif d'une conduite clinique inappropriée [199]. Les logiciels antiplagiat calculent un score de similitude. Un passage identique peut être une citation correcte, avec guillemets et source. Le score n'est pas non plus un indice adéquat de plagiat : il n'existe pas de seuil à partir duquel la similitude signifierait un plagiat, les logiciels signalent aussi des formulations usuelles, et les correspondances doivent être examinées attentivement par un évaluateur [200].

5.6.2 La citation correcte

La citation a deux fonctions : elle reconnaît la contribution d'autrui et permet au lecteur de vérifier l'affirmation. Les revues médicales utilisent le style Vancouver, fondé sur les recommandations du Comité international des rédacteurs de revues médicales (ICMJE) et sur les formats de la Bibliothèque nationale de médecine des États-Unis [201]. Les sources sont numérotées dans l'ordre de leur première apparition dans le texte, et les formats sont mis en pratique lors du TP 5.

5.6.3 L'IA générative dans les travaux des étudiants

Les outils d'IA générative peuvent corriger la langue ou résumer un article, mais ils peuvent aussi inventer des références. Les recommandations de l'ICMJE, mises à jour en janvier 2026, demandent aux auteurs de déclarer, dans la lettre d'accompagnement et dans la section appropriée du travail, comment ils ont utilisé les technologies assistées par l'IA. Les outils de type chatbot (agents conversationnels) ne peuvent pas être auteurs, car ils ne peuvent pas répondre de l'exactitude, de l'intégrité et de l'originalité du travail. Les auteurs doivent pouvoir affirmer que les textes et les images produits avec l'IA ne contiennent pas de plagiat, et le contenu généré par l'IA ne peut pas être cité comme source primaire [202]. Le Comité d'éthique des publications (COPE) a adopté la même position en février 2023 et demande que l'outil et son mode d'utilisation soient décrits dans la section des méthodes [203].

Les politiques des éditeurs ne sont pas uniformes. En 2023, 24 % des 100 plus grands éditeurs et 87 % des 100 revues les mieux classées disposaient de lignes directrices sur l'IA générative. Presque toutes interdisaient l'IA en tant qu'auteur, mais la forme de la déclaration variait beaucoup [204]. Les détecteurs de texte généré n'offrent pas non plus de solution fiable. Ils classent fréquemment comme générés par l'IA des textes écrits par des locuteurs non natifs de l'anglais [205]. Pour l'étudiant, la règle de travail est la même que pour l'auteur d'un article : l'IA peut assister, mais elle ne peut pas signer. Son utilisation se déclare, chaque affirmation se vérifie dans la source originale, et chaque référence se recherche effectivement dans PubMed ou sur le site de l'émetteur.

Pièges fréquents

- "J'ai supprimé le nom et le CNP, donc les données sont anonymes." L'âge, le code postal et un diagnostic rare peuvent identifier le patient, et les données restent personnelles.
- Confondre le consentement éclairé à l'intervention avec la base juridique RGPD des soins. Les soins courants reposent sur l'art. 9, par. 2, point h, et non sur un formulaire de consentement.
- Envoyer des résultats ou des images de patients par des applications de messagerie personnelles. C'est une violation de la confidentialité et, souvent, de la politique de l'hôpital.
- Retarder le signalement d'un incident jusqu'à ce que "la situation soit claire". Le délai de 72 heures court à partir du moment où l'organisation en a eu connaissance.
- Laisser la session ouverte sur l'ordinateur du service ou utiliser le même mot de passe partout.
- Croire qu'un faible score de similitude prouve l'originalité ou que le texte produit par un chatbot n'a pas à être déclaré.

À retenir

- Les données concernant la santé sont une catégorie particulière : leur traitement est interdit, sauf dans les cas prévus à l'art. 9, par. 2, du RGPD.
- Les données pseudonymisées restent des données personnelles. L'anonymisation réelle est difficile à obtenir et doit être démontrée.
- Une violation de données se notifie à l'autorité dans un délai de 72 heures au plus, et les patients sont informés lorsque le risque est élevé.
- La confidentialité, l'intégrité et la disponibilité se protègent ensemble : authentification multifacteur, moindre privilège, journaux d'audit, sauvegardes hors ligne.
- Une cyberattaque contre un hôpital est un événement clinique, et pas seulement technique.
- Le biais algorithmique naît des données et des choix de conception. Les performances se vérifient par sous-groupes.
- Les outils d'IA ne peuvent pas être auteurs. Leur utilisation se déclare, et les sources se vérifient manuellement.

Questions d'autoévaluation

1. (QCS) Selon l'art. 4, point 5, du RGPD, les données pseudonymisées :

a) ne sont plus des données à caractère personnel
b) restent des données à caractère personnel, car elles peuvent être attribuées à la personne à l'aide d'informations supplémentaires
c) sont identiques aux données anonymes
d) peuvent être publiées sans aucune restriction
e) ne peuvent pas être utilisées en recherche

2. (QCS) Un vendredi soir, un hôpital constate qu'un serveur contenant des résultats de laboratoire a été chiffré par un rançongiciel. Le délai dans lequel il notifie, si possible, l'autorité de contrôle est :

a) 24 heures à compter de l'attaque
b) 72 heures à compter du moment où il a pris connaissance de la violation
c) 30 jours à compter de la fin de l'enquête
d) seulement si les patients déposent plainte
e) 72 heures, uniquement les jours ouvrables

3. (QCS) La base juridique RGPD typique du traitement des données d'un patient hospitalisé, à des fins de diagnostic et de traitement, est :

a) le consentement à des fins de marketing
b) l'intérêt légitime de l'hôpital
c) l'art. 9, par. 2, point h, traitement nécessaire à la prise en charge sanitaire, sous l'obligation de secret professionnel
d) les données rendues publiques par le patient
e) aucune base juridique n'est nécessaire

4. (QCM) Parmi les mesures suivantes, lesquelles protègent en premier lieu la confidentialité des données du dossier électronique ?

a) l'authentification multifacteur
b) l'accès par rôle, selon le principe du moindre privilège
c) l'alimentation électrique sans interruption
d) les journaux d'audit des accès, vérifiés périodiquement
e) la redondance des serveurs

5. (QCM) Concernant le biais algorithmique et l'intégrité académique, les affirmations correctes sont :

a) les oxymètres de pouls ont manqué plus souvent l'hypoxémie (hypoxémie occulte) chez les patients noirs, selon les études de 2020-2024
b) un outil d'IA générative peut être mentionné comme coauteur s'il a rédigé une partie du texte
c) un faible score de similitude prouve que le travail ne contient pas de plagiat
d) l'autoplagiat inclut la republication de parties substantielles de ses propres travaux sans citation de l'original
e) les détecteurs de texte généré par l'IA peuvent classer à tort les textes écrits par des locuteurs non natifs de l'anglais

CHAPITRE 6

Chapitre 6. Signaux biomédicaux : acquisition, traitement et interprétation assistée

Objectifs

À la fin de ce chapitre, l'étudiant sera capable :

1. de définir le signal biomédical et de classer les signaux selon leur origine et selon leurs propriétés mathématiques.
2. de décrire la chaîne d'acquisition et de calculer la fréquence d'échantillonnage minimale pour un signal donné.
3. d'expliquer le repliement, la quantification et le rapport signal/bruit sur des exemples numériques simples.
4. de choisir le type de filtre adapté à la dérive de la ligne de base, à l'interférence du réseau à 50 Hz et au bruit de haute fréquence.
5. de calculer la fréquence cardiaque, le SDNN et le RMSSD à partir d'une série d'intervalles RR.
6. d'évaluer de façon critique les algorithmes d'intelligence artificielle pour l'ECG et les dispositifs portables (faux positifs, population de validation, essais randomisés).

6.1 Le signal biomédical

6.1.1 Définition et origine

Un signal biomédical est une grandeur physique variable dans le temps, produite par l'organisme ou obtenue par l'interaction d'un dispositif avec celui-ci, qui porte une information sur une fonction physiologique. L'électrocardiogramme (ECG) suit l'activité électrique du cœur. L'électroencéphalogramme (EEG) suit l'activité électrique du cortex, et l'électromyogramme (EMG) celle des muscles squelettiques.

La classification selon l'origine aide à comprendre quel capteur est nécessaire :

- **bioélectrique** : potentiels générés par des membranes excitables (ECG, EEG, EMG), recueillis à l'aide d'électrodes.
- **mécanique** : pressions, déplacements, forces (pression artérielle invasive, mouvements respiratoires).
- **acoustique** : bruits produits par le cœur, les poumons ou les vaisseaux, recueillis avec un stéthoscope numérique ou un microphone.
- **optique** : variations de la lumière absorbée ou réfléchie par les tissus, comme en photopléthysmographie (*photoplethysmography*, PPG) et en oxymétrie de pouls.
- **chimique** : concentrations de certaines substances mesurées en continu (par exemple le glucose interstitiel) par des capteurs électrochimiques.

La plupart des signaux parviennent à l'ordinateur sous forme de tensions, car un transducteur convertit la grandeur physique (pression, lumière) en tension. L'oxymètre de pouls, par exemple, envoie une lumière rouge et infrarouge à travers les tissus, et une photodiode enregistre la variation d'intensité à chaque pulsation artérielle. À partir de ces variations, après étalonnage, l'appareil estime la saturation en oxygène (SpO2) [206].

6.1.2 Classifications utiles

Le signal **analogique** peut prendre n'importe quelle valeur dans un intervalle continu, tandis que le signal **numérique** ne prend que des valeurs discrètes dans un domaine défini [62]. De plus, le signal numérique n'est connu qu'aux instants d'échantillonnage. La tension présente sur la peau du thorax est analogique. Le fichier enregistré par l'électrocardiographe est numérique.

La deuxième classification concerne la prévisibilité. Un signal **déterministe** peut être entièrement décrit par une formule (une sinusoïde de 50 Hz produite par le réseau électrique). Un signal **aléatoire** (stochastique) ne peut être décrit que statistiquement, par sa moyenne, sa variance et son spectre. Les signaux physiologiques réels sont des mélanges : l'ECG d'un adulte au repos paraît presque **périodique**, avec une forme d'onde qui se répète à chaque battement. Pourtant, l'intervalle entre les battements varie continuellement, et cette variation a elle-même une valeur clinique (section 6.3.6). Les signaux biologiques sont donc, tout au plus, quasi périodiques.

6.1.3 Exemples et valeurs typiques

Le tableau 6.1 réunit les signaux que l'on rencontre le plus souvent en clinique. Les valeurs sont indicatives.

Tableau 6.1. Caractéristiques de quelques signaux biomédicaux usuels. Source : [206-211].

Signal	Origine	Amplitude ou unité	Bande ou réglages usuels
ECG	bioélectrique (myocarde)	environ 1 mV sur la peau	passe-haut 0,05 Hz pour l'ECG standard, 0,5 Hz ou plus sur les moniteurs
EEG de scalp	bioélectrique (cortex)	microvolts (rythme bêta : 10-20 µV)	affichage 0,5-70 Hz, échantillonnage minimal 256 Hz
EMG de surface	bioélectrique (unités motrices)	variable selon le muscle et l'effort	passe-bande 20-500 Hz, énergie dominante 20-150 Hz
PPG	optique (volume sanguin pulsatile)	unités relatives	composante principale 0,5-5 Hz
Pression artérielle invasive	mécanique (cathéter et transducteur)	mmHg	composantes pulsatiles jusqu'à 20 Hz
SpO2	optique (rouge et infrarouge)	pourcentage de saturation	valeur dérivée de la composante pulsatile

Le signal ECG recueilli sur la peau n'atteint qu'environ 1 mV, c'est pourquoi il est sensible aux artefacts produits par les mouvements du patient ou des électrodes et à l'interférence du réseau électrique. Sa qualité s'est améliorée grâce à des techniques analogiques et numériques de traitement [206]. Pour l'EEG, dont les amplitudes sont de l'ordre du microvolt, le problème est encore plus important. En électroencéphalographie, le clinicien travaille avec des bandes de fréquence désignées par des lettres grecques (tableau 6.2). Le rythme alpha postérieur, de 8-12 Hz, définit l'activité de fond normale de l'EEG de l'adulte éveillé. Il est plus net les yeux fermés et s'atténue à l'ouverture des yeux [209].

Tableau 6.2. Bandes de fréquence de l'EEG et contexte physiologique typique. Source : [208,209].

Bande	Fréquence	Contexte physiologique typique
Delta (δ)	0,5-4 Hz	sommeil profond
Thêta (θ)	4-7 Hz	somnolence, stades N1 et N2
Alpha (α)	8-12 Hz	veille détendue, yeux fermés, région occipitale
Sigma (σ)	12-16 Hz	fuseaux de sommeil au stade N2
Bêta (β)	13-30 Hz	veille, régions frontale et centrale
Gamma (γ)	30-80 Hz	traitement sensoriel et cognition

6.2 La chaîne d'acquisition

6.2.1 De l'électrode au convertisseur

La figure 6.1 montre le trajet d'un signal depuis la source biologique jusqu'à l'interprétation. Le premier élément est le **capteur** (l'électrode, le transducteur de pression, la photodiode). Vient ensuite un **amplificateur différentiel**, qui amplifie la différence de tension entre deux électrodes et rejette les tensions communes aux deux, comme celle qu'induit le réseau électrique. La capacité de réjection dépend de l'équilibre des impédances des électrodes : le standard minimal IFCN-ILAE pour l'EEG recommande des impédances inférieures à 5 kΩ et considère comme acceptables des valeurs inférieures à 10 kΩ [208].

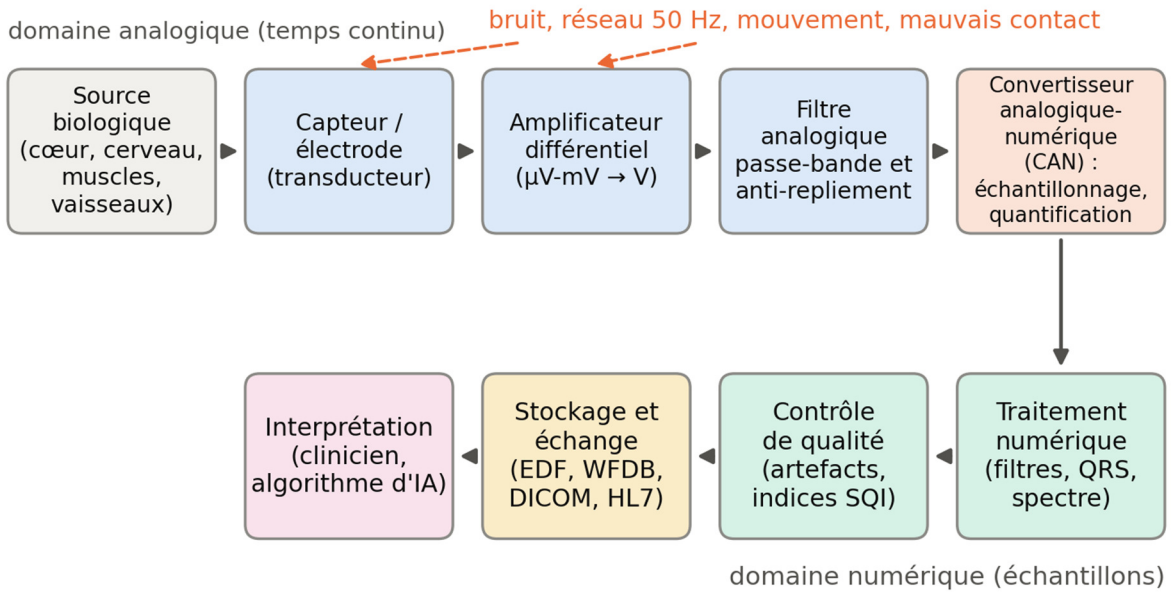


Figure 6.1. Chaîne d'acquisition et de traitement d'un signal biomédical. La rangée du haut travaille sur des tensions continues dans le temps, la rangée du bas sur des échantillons numériques. Schéma original.

Le **filtre analogique** placé en amont du convertisseur a deux rôles. Il élimine les composantes dépourvues de valeur diagnostique (la dérive lente, le bruit de très haute fréquence) et limite la bande du signal en dessous de la moitié de la fréquence d'échantillonnage, rôle que l'on appelle anti-repliement (*anti-aliasing*). Le dernier bloc analogique est le **convertisseur analogique-numérique** (CAN, en anglais *analog-to-digital converter*, ADC), qui transforme les valeurs échantillonnées d'un signal continu en une représentation numérique discrète [62].

Pour la pression artérielle invasive, la chaîne comprend le cathéter, la tubulure remplie de sérum physiologique et le transducteur. Deux sources d'erreur dépendent du personnel : la mise à zéro au niveau du cœur et la réponse dynamique, vérifiée par le test de rinçage rapide (*fast flush test*), car de petites bulles d'air déforment l'onde [206].

6.2.2 L'échantillonnage et le théorème de Nyquist-Shannon

L'échantillonnage (*sampling*) consiste à mesurer la valeur du signal analogique à intervalles de temps égaux. Le nombre de mesures par seconde est la fréquence d'échantillonnage, fe, exprimée en hertz. Le théorème d'échantillonnage (Nyquist-Shannon) énonce qu'un signal qui ne contient aucune composante au-delà de la fréquence fmax peut être reconstitué sans perte si fe > 2·fmax. Le manuel de Shortliffe et Cimino appelle ce seuil minimal fréquence de Nyquist [62]. Autrement dit, la plus haute fréquence correctement restituée est fe/2 : à 256 Hz, 128 Hz [208]. Certains textes appellent fréquence de Nyquist la valeur fe/2 elle-même, si bien que le sens du terme doit être vérifié dans chaque source.

Si la règle n'est pas respectée, il se produit un **repliement** (*aliasing*) : une composante de haute fréquence se "déguise" en une fréquence fausse, plus basse. La figure 6.2 montre le phénomène sur une sinusoïde de 5 Hz. À 50 échantillons par seconde, sa forme est conservée. À 6 échantillons par seconde, les points décrivent une onde de 1 Hz, qui n'existe pas dans le signal réel.

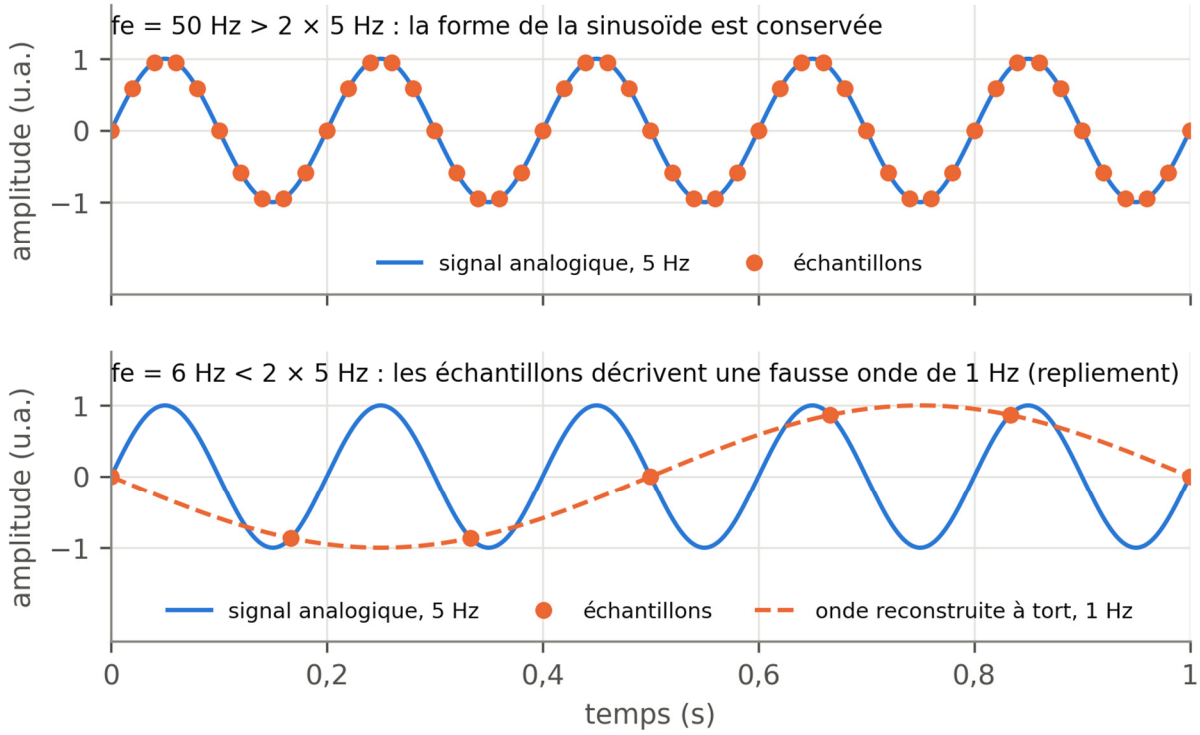


Figure 6.2. Échantillonnage correct (en haut) et repliement (en bas) pour une sinusoïde de 5 Hz. La fréquence apparente en cas d'échantillonnage insuffisant est |5 - 6| = 1 Hz. Graphique calculé, données synthétiques, illustratives.

Exemple

- **Fréquence minimale pour l'EEG.** Les composantes cliniques de l'EEG de routine sont affichées jusqu'à 70 Hz. Le théorème exige fe > 140 Hz. En raison de l'alignement de phase, le standard IFCN-ILAE recommande un échantillonnage à au moins 2,5 fois la fréquence maximale et propose un minimum de 256 Hz [208].
- **Repliement du réseau électrique.** Une interférence de 50 Hz échantillonnée à 60 Hz sans filtre anti-repliement apparaît à |50 - 60| = 10 Hz, c'est-à-dire exactement dans la bande alpha de l'EEG. Un artefact technique pourrait être confondu avec un rythme cérébral.
- **Précision temporelle.** À 250 Hz, l'instant de l'onde R est localisé avec une erreur d'environ ±4 ms. À 1 000 Hz, l'erreur descend à ±1 ms, et c'est pourquoi 1 000 Hz est la fréquence recommandée, surtout lorsque l'onde R doit être mise en relation avec d'autres événements synchronisés [212].

6.2.3 La quantification et la résolution

Après l'échantillonnage, chaque valeur doit être écrite avec un nombre fini de bits. Ce processus s'appelle la quantification (*quantization*). Un CAN de N bits divise le domaine d'entrée en 2^N niveaux. Le pas de quantification (la résolution) est égal au domaine divisé par le nombre de niveaux.

L'exemple classique provient de la base MIT-BIH Arrhythmia, disponible sur PhysioNet : les signaux ont été numérisés à 360 échantillons par seconde et par canal, avec une résolution de 11 bits sur un domaine de 10 mV [213]. On obtient 2^11 = 2 048 niveaux et un pas de 10 mV / 2 048 ≈ 4,9 µV. Un complexe QRS de 1 mV couvre ainsi environ 200 niveaux. Sur 16 bits, le même domaine compterait 65 536 niveaux, avec un pas d'environ 0,15 µV.

La quantification ajoute une petite erreur, d'au plus un demi-pas. Pour un signal qui occupe tout le domaine, le rapport théorique entre le signal et le bruit de quantification est d'environ 6,02·N + 1,76 dB. À 11 bits, on obtient environ 68 dB, bien au-dessus du bruit biologique et de celui des électrodes.

Le volume des données est le produit du nombre de canaux, de la fréquence d'échantillonnage et de la durée. Un enregistrement Holter de 24 heures, à 2 canaux et à 360 Hz, compte 2 × 360 × 86 400 ≈ 62 millions d'échantillons, soit environ 124 Mo non compressés à 2 octets par échantillon.

6.2.4 Le rapport signal/bruit

Le rapport signal/bruit (*signal-to-noise ratio*, SNR) compare la puissance du signal utile à celle du bruit : SNR = 10·log10(Ps/Pb), en décibels. Lorsqu'on compare des amplitudes, la formule devient 20·log10(As/Ab). Un ECG d'une amplitude de 1 mV superposé à un bruit de 0,1 mV a un SNR de 20·log10(10) = 20 dB. Avec un bruit de 0,5 mV, le SNR tombe à environ 6 dB. Le SNR estimé sur de courtes fenêtres de signal est l'un des indices statistiques utilisés pour décider automatiquement si un segment d'ECG est exploitable [214].

6.3 Le traitement numérique

6.3.1 Les filtres

Un filtre numérique modifie le spectre du signal : il laisse passer certaines fréquences et atténue les autres. En pratique clinique, on rencontre quatre types de filtres :

- **passe-bas** : conserve les basses fréquences et atténue le bruit de haute fréquence (par exemple l'activité musculaire superposée à l'ECG).
- **passe-haut** : élimine les composantes très lentes, comme la dérive de la ligne de base due à la respiration, à la transpiration ou au mouvement des électrodes.
- **passe-bande** : combinaison des deux, utilisée pour l'EMG de surface dans l'intervalle typique de 20-500 Hz [210].
- **coupe-bande** (*notch filter*) : élimine une bande étroite autour d'une fréquence fixe, généralement la fréquence du réseau électrique, 50 Hz en Europe et 60 Hz en Amérique du Nord [210].

Le filtrage a un prix. Un filtre passe-haut agressif peut déformer le segment ST de l'ECG. L'ECG standard à 12 dérivations utilise un filtre passe-haut de 0,05 Hz, alors que les moniteurs de chevet utilisent souvent 0,5 Hz ou plus pour réduire la dérive. Dans une analyse de 2026 portant sur 150 patients, le filtre de 1,0 Hz a multiplié par près de dix la cote d'un faux sus-décalage du segment ST d'au moins 1 mm en V1 (OR 9,14) [207]. Règle pratique : un ST suspect sur le moniteur se vérifie sur un ECG standard à 12 dérivations. Même les manuels d'informatique médicale notent que le filtrage nécessaire à la surveillance continue "déforme légèrement" l'ECG [206].

La figure 6.3 montre l'effet d'une chaîne simple de filtres sur un ECG synthétique. Après filtrage, les pics R peuvent être détectés à l'aide d'un simple seuil.

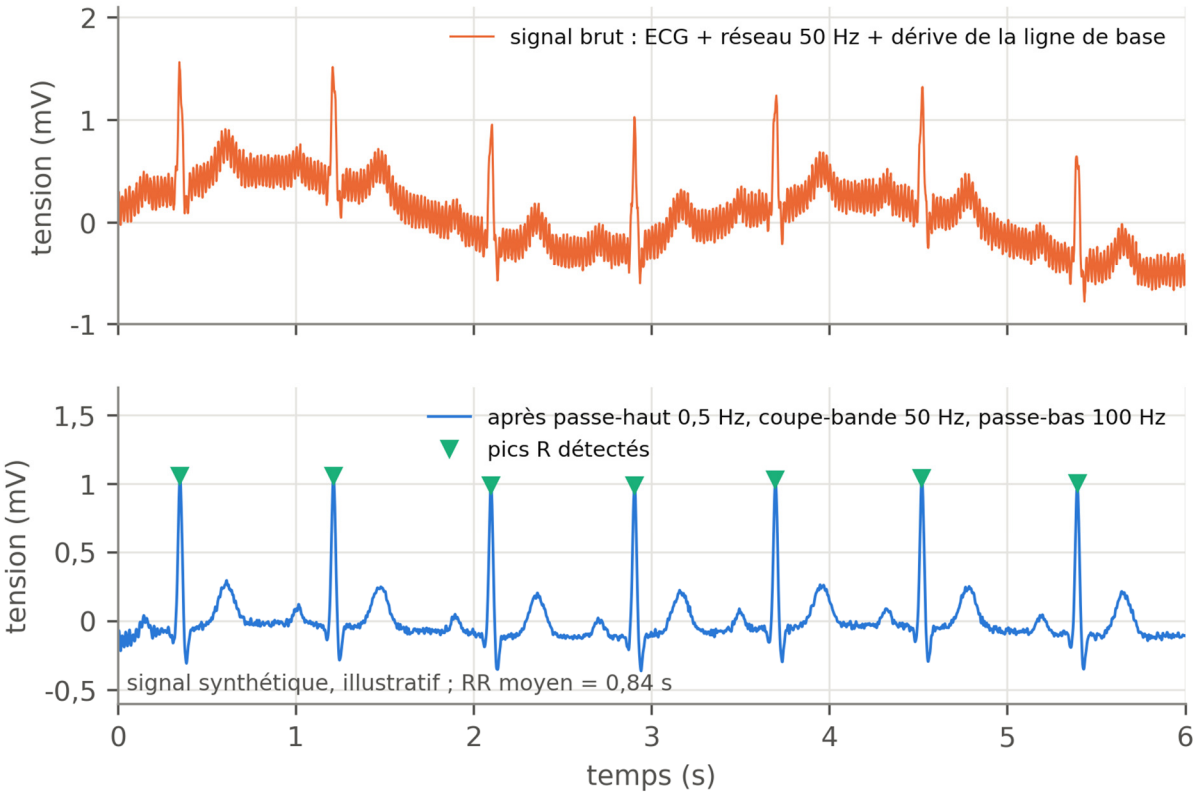


Figure 6.3. ECG synthétique avec interférence de 50 Hz et dérive de la ligne de base (en haut), puis après les filtres passe-haut 0,5 Hz, coupe-bande 50 Hz et passe-bas 100 Hz (en bas), avec les pics R détectés automatiquement. Signal synthétique, illustratif, généré à partir d'ondes gaussiennes.

6.3.2 Le moyennage

Certains signaux sont beaucoup plus faibles que le bruit de fond, mais se répètent à l'identique après un stimulus : les potentiels évoqués visuels ou auditifs en sont l'exemple typique. Si l'on

aligne N réponses sur l'instant du stimulus et que l'on en calcule la moyenne, la composante synchronisée est conservée, tandis que le bruit aléatoire est divisé par √N. Avec 100 répétitions, le bruit est divisé par 10, autrement dit le SNR augmente de 20 dB. La méthode suppose une réponse identique à chaque répétition.

6.3.3 La transformée de Fourier et le spectre

Tout signal peut être décomposé en une somme de sinusoïdes de fréquences différentes. La transformée de Fourier discrète, calculée efficacement par l'algorithme FFT (*fast Fourier transform*), fait passer le signal du domaine temporel au domaine fréquentiel. Le résultat, le spectre de puissance, montre quelle quantité d'énergie le signal possède à chaque fréquence. Les recommandations de l'IFCN pour l'EEG de repos comptent précisément le filtrage par bandes et le calcul du spectre d'amplitude ou de puissance parmi les étapes de base de l'extraction des caractéristiques [215].

Sur le spectre, l'interférence du réseau apparaît comme un pic étroit à 50 Hz, et la dérive de la ligne de base comme une énergie concentrée sous 0,5 Hz (figure 6.4). La même figure montre que la plus grande partie de l'énergie de l'ECG synthétique se situe sous 40 Hz.

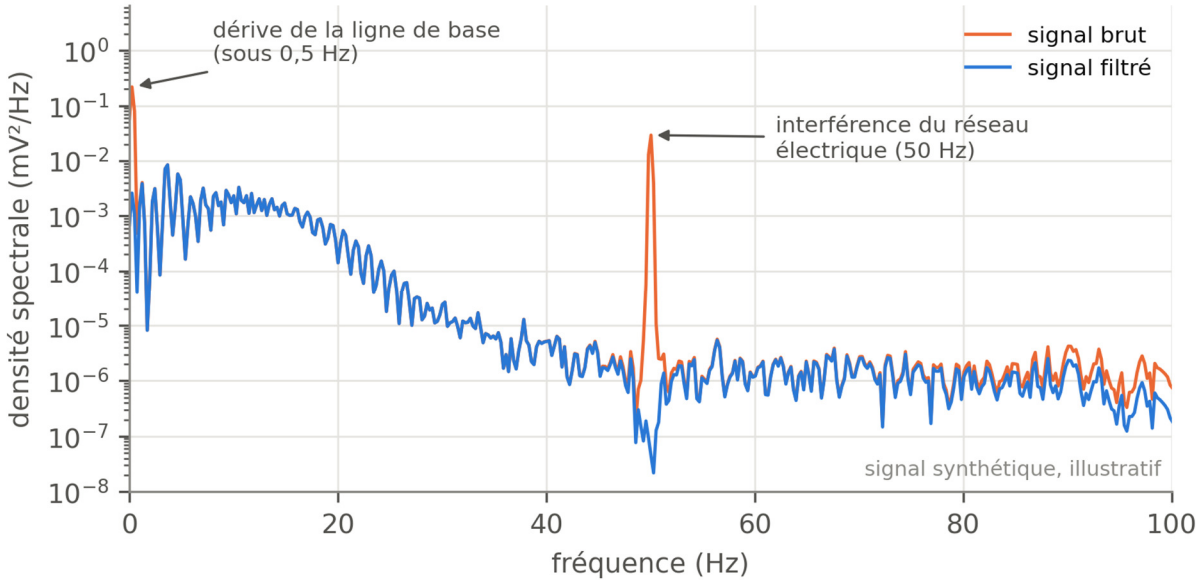


Figure 6.4. Densité spectrale de puissance (méthode de Welch) de l'ECG synthétique de la figure 6.3, avant et après filtrage. Le pic de 50 Hz et l'énergie située sous 0,5 Hz disparaissent après filtrage. Signal synthétique, illustratif.

6.3.4 L'analyse temps-fréquence, en bref

Le spectre classique suppose un signal aux propriétés constantes, or une crise d'épilepsie commence brusquement et un fuseau de sommeil dure une seconde. L'analyse temps-fréquence montre "quelles fréquences sont présentes, et à quel moment". Le spectrogramme applique la transformée de Fourier sur de courtes fenêtres qui glissent dans le temps, tandis que la transformée en ondelettes (*wavelet*) utilise des fenêtres courtes pour les hautes fréquences et longues pour les basses. La résolution temporelle et la résolution fréquentielle ne peuvent pas être simultanément aussi fines qu'on le souhaiterait.

6.3.5 La détection des complexes QRS

Le détecteur de QRS constitue la première étape de presque toute analyse automatique du rythme. Un détecteur classique, du type Pan-Tompkins, procède en quelques étapes : il applique au signal un filtre passe-bande pour conserver l'énergie du complexe QRS, calcule la dérivée du

signal pour faire ressortir les pentes abruptes, élève les valeurs au carré, les intègre sur une fenêtre glissante de l'ordre de 150 ms et compare le résultat à un seuil adaptatif. Des logiciels de visualisation gratuits, comme EDFbrowser, intègrent un tel détecteur et permettent d'exporter les intervalles RR [216].

Les détecteurs ne sont pas interchangeables. Une étude de 2026 a comparé huit algorithmes, dont Pan-Tompkins, sur des enregistrements de 25 volontaires, au repos, sous effort cognitif et pendant une activité physique. Aucun algorithme n'a été le meilleur dans toutes les conditions, et l'accord des indicateurs de variabilité de la fréquence cardiaque avec les valeurs de référence a varié de presque parfait à nul, selon le détecteur et le mode d'acquisition [217].

6.3.6 La variabilité de la fréquence cardiaque

L'intervalle RR est le temps qui sépare deux pics R successifs. La fréquence cardiaque instantanée est FC = 60/RR (RR en secondes) : un RR de 0,8 s correspond à une FC de 75 battements par minute. La variabilité de la fréquence cardiaque (VFC, en anglais *heart rate variability*, HRV) décrit les fluctuations de ces intervalles d'un battement à l'autre. Elle se calcule uniquement sur les intervalles NN, c'est-à-dire entre des battements sinusaux normaux. Les battements ectopiques et les artefacts doivent être identifiés et traités avant le calcul, de préférence par une inspection visuelle après la détection automatique, car les valeurs signalées automatiquement comme aberrantes ne sont pas toujours des artefacts, et un battement ectopique ne peut pas être simplement "corrigé" [212].

Tableau 6.3. Indicateurs de base de la variabilité de la fréquence cardiaque. Source : [212].

Indicateur	Définition	Remarque
SDNN (ms)	écart-type des intervalles NN	variabilité globale, dépendante de la durée de l'enregistrement
RMSSD (ms)	racine carrée de la moyenne des carrés des différences successives	variabilité à court terme, abolie par le blocage vagal
pNN50 (%)	pourcentage des différences successives supérieures à 50 ms	apparenté au RMSSD
HF	puissance spectrale dans la bande respiratoire, typiquement 0,12-0,40 Hz	reflète l'arythmie sinusale respiratoire
LF	puissance spectrale dans la bande 0,05-0,15 Hz (parfois à partir de 0,04 Hz)	interprétation physiologique controversée

Exemple

- Cinq intervalles NN consécutifs : 800, 820, 790, 810 et 830 ms. La moyenne est de 810 ms, donc FC moyenne = 60 000/810 ≈ 74 bpm.
- **SDNN** : les écarts à la moyenne sont de -10, 10, -20, 0 et 20 ms. La somme des carrés vaut 1 000, qui, divisée par n - 1 = 4, donne 250, et la racine carrée donne SDNN ≈ 15,8 ms.
- **RMSSD** : les différences successives sont de 20, -30, 20 et 20 ms. Leurs carrés (400, 900, 400, 400) ont pour moyenne 525, et la racine carrée donne RMSSD ≈ 22,9 ms.
- Cinq intervalles sont trop peu nombreux pour tirer une conclusion clinique. L'exemple montre seulement la mécanique du calcul, reprise dans le TP 6.

Pour l'analyse spectrale, la série RR, non uniforme dans le temps, est rééchantillonnée à intervalles égaux, de préférence à 4 Hz [212]. L'interprétation exige de la prudence : la puissance LF n'est pas un indice simple de l'activité sympathique, et la bande HF standard peut manquer la respiration très lente d'un adulte ou la respiration plus rapide des enfants [212].

6.4 Artefacts et qualité du signal

6.4.1 Types d'artefacts

Un artefact est toute composante du signal qui ne provient pas du phénomène physiologique étudié. Pour l'ECG, les sources décrites dans la littérature comprennent les électrodes décollées, le mouvement, les fils rompus, les dispositifs externes ou implantés et même d'autres signaux physiologiques [218]. Les artefacts de l'ECG peuvent imiter une tachycardie ventriculaire ou un infarctus aigu du myocarde, et leur reconnaissance relève donc de la compétence clinique du médecin [218].

Les principales catégories sont les suivantes :

- **le mouvement** : le déplacement de l'électrode par rapport à la peau modifie le potentiel de contact et produit des oscillations lentes ou brusques. En EMG, les artefacts de mouvement se situent typiquement entre 0 et 20 Hz [210].
- **le contact de l'électrode** : gel desséché, impédance élevée ou déséquilibrée, électrode qui "saute" (*electrode pop*). Les électrodes à impédance élevée sont davantage perturbées par la transpiration et le mouvement [208].
- **le réseau électrique** : une composante sinusoïdale étroite à 50 ou 60 Hz, induite par les câbles et les appareils alimentés par le secteur [210].
- **l'activité musculaire (EMG)** superposée à l'ECG ou à l'EEG : bruit de haute fréquence, fréquent chez les patients agités, qui frissonnent ou présentent un tremblement.
- **les artefacts physiologiques en EEG** : les mouvements oculaires, le pouls, l'ECG et la transpiration. Certains, comme l'EMG et les mouvements oculaires, aident même à l'interprétation [219].

En PPG, l'artefact de mouvement (environ 0,01-10 Hz) se superpose à la bande principale du signal, 0,5-5 Hz, et peut la déformer, si bien qu'il doit être éliminé ou corrigé avant l'analyse [211]. L'oxymètre de pouls présente des limites similaires : le mouvement, l'anémie, une faible perfusion périphérique et la peau froide. Il figure parmi les dispositifs de soins intensifs qui produisent le plus de fausses alarmes [206].

6.4.2 Le contrôle de qualité

Le contrôle de qualité commence au chevet du patient, avec la préparation de la peau et la fixation des électrodes et des câbles. En EEG, on vérifie l'impédance de chaque électrode au début de l'enregistrement [208]. Pour la pression invasive, la mise à zéro et le test de rinçage rapide dépendent du personnel, et le système cathéter-transducteur exige une surveillance continue : un prélèvement sanguin par le cathéter, par exemple, fait disparaître temporairement l'onde pulsatile [206].

Le contrôle se poursuit dans le logiciel, au moyen d'indices de qualité du signal (*signal quality indices*, SQI) : la puissance relative du complexe QRS, l'asymétrie (*skewness*), l'aplatissement (kurtosis), le SNR. Une analyse de 2022 a montré que ces indices statistiques, utilisés avec des seuils fixes, se comportent différemment d'une base de données à l'autre, et les auteurs

recommandent donc des seuils adaptatifs ou des méthodes d'apprentissage automatique [214]. Une revue systématique des techniques automatiques d'évaluation de la qualité de l'ECG publiées en 2012-2022 a rapporté une exactitude comprise entre 85,75 % et 97,15 %, mais aussi une quasi-absence de tests en temps réel chez l'humain [220].

L'enjeu clinique est la fatigue face aux alarmes. Les données citées dans le manuel de Shortliffe et Cimino font état de 1,6 à 14,6 alarmes par patient et par heure en soins intensifs, dont jusqu'à 90 % de fausses alarmes. Quelques secondes d'artefact sur l'ECG peuvent déclencher, par exemple, une fausse alarme de tachycardie ventriculaire [206]. Parmi les solutions proposées figurent de courts délais avant le déclenchement de l'alarme et la combinaison de plusieurs signaux qui estiment la même fréquence cardiaque [206].

6.5 Formats et bases de données ouvertes

6.5.1 Formats de fichier

Sans métadonnées (fréquence d'échantillonnage, unité, gain, dérivation), les échantillons sont presque inutilisables. Les formats standard les conservent ensemble.

- **EDF et EDF+**. Le format européen de données (*European Data Format*) a été publié en 1992, et EDF+ en 2003. EDF+ peut contenir des enregistrements discontinus, des annotations, des stimuli et des événements. Le format est surtout utilisé en neurophysiologie clinique, en somnologie et en cardiologie, et il est devenu le standard de facto pour l'EEG et la polysomnographie [221]. Le standard minimal IFCN-ILAE recommande que les données EEG puissent être exportées en CSV, en EDF ou en DICOM [208].
- **WFDB** (*Waveform Database*) est un ensemble de standards ouverts et de logiciels destinés au stockage et à l'analyse des signaux physiologiques et des annotations, développé au MIT à partir de 1989. Un enregistrement WFDB comprend un fichier d'en-tête (.hea), un ou plusieurs fichiers de signal (.dat) et des fichiers d'annotations, et ce format est à la base de nombreuses bases de données de PhysioNet [222].
- **DICOM waveform**. Le standard DICOM, connu en imagerie, définit aussi des objets pour les formes d'onde : ECG à 12 dérivations, ECG général et ambulatoire, formes d'onde hémodynamiques, pouls artériel, respiration, EEG de scalp de routine, EEG de sommeil, EMG et électro-oculogramme [223].
- **HL7 aECG** (*annotated ECG*) est un standard XML créé en réponse à l'initiative de la FDA sur l'ECG numérique de 2001. Son objectif initial était l'évaluation systématique des formes d'onde et des points de mesure (par exemple l'intervalle QT) dans les essais cliniques de médicaments non cardiologiques [224].
- **ISO/IEEE 11073** est la famille de normes pour la communication des dispositifs médicaux au point de soins. La variante historique destinée au chevet du patient (Medical Information Bus) a connu une faible adoption commerciale [206]. La partie 10103 (édition publiée en 2025) définit les termes du résumé d'interrogation des dispositifs cardiaques implantables (stimulateurs, défibrillateurs, dispositifs de resynchronisation, moniteurs implantables), transféré des systèmes des fabricants vers le dossier de santé électronique [225].

6.5.2 PhysioNet, une ressource ouverte

PhysioNet, maintenu par le Laboratoire de physiologie computationnelle du MIT, héberge des bases de données de signaux physiologiques, des logiciels ouverts et des compétitions scientifiques (*PhysioNet Challenges*) [226]. En 2026, l'équipe de la ressource l'a décrite comme

une plateforme mondiale pour la recherche biomédicale, et cet article est la citation standard demandée aux utilisateurs [227].

La base MIT-BIH Arrhythmia contient 48 extraits d'une demi-heure chacun, provenant de 47 sujets, avec environ 110 000 annotations de battements vérifiées indépendamment par au moins deux cardiologues. L'accès est libre, sous licence Open Data Commons Attribution [213]. Le TP 6 utilise cette base et le visualiseur LightWAVE, qui fonctionne directement dans le navigateur [228].

6.6 L'interprétation assistée par l'intelligence artificielle

6.6.1 L'ECG et l'intelligence artificielle

L'interprétation automatique de l'ECG à 12 dérivations est utilisée depuis les années 1980. Les algorithmes sont assez précis, surtout chez les personnes en bonne santé, mais les désaccords avec les cardiologues peuvent être cliniquement importants, c'est pourquoi le cardiologue confirme habituellement le résultat [206]. Les modèles d'intelligence artificielle (IA) fondés sur l'apprentissage profond (*deep learning*) ont élargi les questions posées à l'ECG : dysfonction systolique du ventricule gauche, cardiomyopathie hypertrophique, amylose cardiaque, sténose aortique ou risque de fibrillation atriale future. Une revue de 2024 demande, avant une adoption à large échelle, des preuves de bénéfice pour les patients et des mises en œuvre qui ne submergent pas le système de résultats faux positifs [229].

L'exemple de référence est EAGLE, un essai pragmatique randomisé en grappes, dans lequel 120 équipes de soins primaires de 45 cliniques ou hôpitaux ont reçu ou non le résultat d'un algorithme ECG qui signalait la probabilité d'une fraction d'éjection basse (≤ 50 %). Au total, 22 641 adultes sans insuffisance cardiaque connue ont été inclus. Dans le bras ayant accès à l'IA, les nouveaux diagnostics de fraction d'éjection basse à 90 jours sont passés de 1,6 % à 2,1 % (OR 1,32). Parmi les patients dont le résultat de l'IA était positif (6 % de la cohorte), ces diagnostics sont passés de 14,5 % à 19,5 %. Le nombre total d'échocardiographies est resté similaire (19,2 % contre 18,2 %) [230].

L'idée a aussi été testée dans d'autres systèmes de santé. Un essai randomisé pragmatique portant sur 13 631 patients hospitalisés à Taïwan, pris en charge par des médecins non cardiologues, a relevé davantage de nouveaux diagnostics de fraction d'éjection basse à 30 jours (1,5 % contre 1,1 %, HR 1,50). Le recours à l'échocardiographie n'a pas augmenté, mais les échocardiographies demandées ont plus souvent confirmé la maladie (34,2 % contre 20,2 %) [231]. Au Nigeria, un essai randomisé portant sur 1 232 femmes enceintes ou en post-partum a comparé un dépistage guidé par l'IA aux soins habituels. Le stéthoscope numérique doté d'IA a doublé la détection de la dysfonction systolique (4,1 % contre 2,0 %, OR 2,12). Le modèle appliqué à l'ECG à 12 dérivations a eu en revanche un effet de même sens qui n'a pas atteint le seuil de signification statistique (3,4 % contre 2,0 %, p = 0,125) [232]. Cette différence suggère qu'un modèle performant dans une population n'apporte pas automatiquement le même bénéfice dans une autre.

L'IA peut aussi extraire de l'ECG des informations invisibles à l'œil nu. Chez 1 936 personnes âgées sans fibrillation atriale connue, la sortie d'un modèle appliqué à un ECG en rythme sinusal a prédit la fibrillation atriale future avec une statistique C de 0,69, égale à celle du score clinique CHARGE-AF. La combinaison des deux a atteint 0,72 [233]. Une statistique C de 0,69 est modeste : le modèle hiérarchise le risque, il ne pose pas le diagnostic.

6.6.2 La fibrillation atriale détectée par les montres et les bracelets

Les montres connectées utilisent deux technologies. La PPG optique mesure le pouls en continu et signale un rythme irrégulier. L'ECG à une seule dérivation s'enregistre à la demande, le doigt posé sur le boîtier de la montre. Les capteurs PPG suivent déjà la fréquence et le rythme cardiaques, le sommeil et l'effort, et la feuille de route du domaine publiée en 2023 décrit les recherches nécessaires à leur utilisation clinique [234].

Une étude de grande envergure, la Fitbit Heart Study, a inclus 455 699 participants aux États-Unis. Des notifications de rythme irrégulier sont apparues chez 1 % d'entre eux et chez 4 % des participants âgés de 65 ans ou plus. Chez 1 057 d'entre eux, un patch ECG porté pendant une semaine a confirmé la fibrillation atriale dans 32,2 % des cas. Lorsque la notification réapparaissait pendant la surveillance ECG, la valeur prédictive positive était de 98,2 % [235]. Une analyse ultérieure de la même cohorte a montré que la probabilité de confirmation dépend de la durée de la surveillance : 7,6 % pour un seul ECG de 10 secondes, 32,2 % pour un moniteur porté une semaine et 60,8 % pour un moniteur porté quatre semaines [236].

Un essai randomisé publié en 2026 avec l'Apple Watch (EQUAL) a inclus 437 patients âgés d'au moins 65 ans, à risque élevé d'accident vasculaire cérébral. Après six mois de surveillance par la montre, la fibrillation atriale nouvellement diagnostiquée était plus fréquente que dans le groupe recevant les soins standard (9,6 % contre 2,3 %, HR 4,40) [237]. Les méta-analyses d'exactitude donnent des chiffres élevés. L'une d'elles, publiée en 2025 et portant sur 26 études et 17 349 patients, a trouvé une sensibilité de 95 % et une spécificité de 97 %, comparables entre les modèles PPG et ECG, mais avec une hétérogénéité très importante entre les études (I^2 de 95 % et 99 %) [238].

6.6.3 Limites

Une sensibilité et une spécificité élevées ne signifient pas automatiquement un bénéfice clinique. La collaboration AF-SCREEN note que les essais randomisés de dépistage menés par des professionnels n'ont pas démontré de réduction des accidents vasculaires cérébraux. La suspicion issue de la PPG se confirme par un ECG, et l'ECG enregistré par le consommateur est lu par un professionnel. Chez les jeunes qui adoptent les premiers la technologie, il est peu probable que le dépistage soit coût-efficace [239]. Une analyse de 2026, explicitement génératrice d'hypothèses, a synthétisé les essais randomisés de surveillance par dispositifs portables chez des personnes asymptomatiques à risque élevé. La surveillance continue a augmenté la cote d'un diagnostic de fibrillation atriale (OR 2,84) et doublé le recours aux anticoagulants, sans réduction nette des événements indésirables, ce qui suggère un surdiagnostic [240].

La deuxième limite est la population. Un modèle entraîné sur une population homogène peut se comporter différemment dans une autre. Pour le modèle ECG de détection de la fraction d'éjection basse, dérivé d'une cohorte composée à 96,2 % de personnes blanches non hispaniques, la performance est restée constante entre les groupes raciaux et ethniques (AUC 0,93-0,96). Les auteurs recommandent néanmoins de rapporter la performance par sous-groupes d'âge, de sexe, de race et d'ethnie pour tout nouvel outil [241]. Pour les capteurs optiques, le problème est physique : la mélanine absorbe la lumière. Une méta-analyse de 2024 a montré que les oxymètres de pouls surestiment la SpO2 aussi bien sur les peaux claires que sur les peaux foncées, même si cette surestimation peut ne pas être cliniquement pertinente. La fréquence du pouls mesurée par les dispositifs portables ne présentait pas d'erreur systématique liée à la pigmentation [242].

La troisième limite est la validation externe : les résultats obtenus dans le centre qui a développé le modèle doivent être reproduits sur d'autres appareils, dans d'autres hôpitaux et dans d'autres populations. L'étude nigériane a appliqué des modèles développés à la Mayo Clinic, entraînés sur des données indépendantes de celles de l'étude [232]. L'étude taïwanaise a en revanche testé un modèle développé auparavant par la même équipe : elle montre donc le bénéfice dans un autre contexte clinique, et non une validation externe proprement dite [231].

La surveillance continue ne se limite pas à la cardiologie. Une revue systématique de l'auteur a inclus 9 études (778 patients) utilisant des capteurs portables après une chirurgie colorectale. Elle a constaté que le nombre de pas des premiers jours postopératoires était associé à la durée d'hospitalisation, aux complications et à la réhospitalisation, tout en appelant à des définitions standardisées et à une validation multicentrique [243].

6.7 La situation en Roumanie en 2026

Au 27 septembre 2026, les bases de données ouvertes de PhysioNet, les formats EDF et WFDB et les logiciels gratuits décrits dans ce chapitre peuvent aussi être utilisés dans les universités roumaines, sous des licences ouvertes qui exigent la citation de la source. Du côté des systèmes, le cahier des charges publié par le ministère de la Santé (Ministerul Sănătății) pour le système informatique intégré des services d'anesthésie-réanimation (SIEC-ATI) concerne 18 hôpitaux, 649 postes-lits de soins intensifs et 275 postes d'anesthésie. Le document exige l'enregistrement automatique, au plus toutes les 90 secondes, des paramètres mesurés par les appareils connectés au patient (moniteur de fonctions vitales, ventilateur, pompes), l'émission automatique d'alarmes et des interfaces HL7/FHIR [244]. L'exigence porte sur des valeurs numériques périodiques (tendances), et non sur les formes d'onde échantillonnées à des centaines de hertz, différence que l'étudiant peut expliquer après avoir étudié la section 6.2. Le chapitre n'évalue pas l'état d'avancement de la mise en œuvre du système.

Pièges fréquents

- La confusion entre la fréquence du signal et la fréquence d'échantillonnage : un signal dont les composantes vont jusqu'à 70 Hz exige plus de 140 Hz.
- La croyance que le repliement se corrige après la numérisation. La fréquence fausse ne peut plus être séparée, et c'est pourquoi le filtre anti-repliement est placé avant le convertisseur.
- L'interprétation d'un sus-décalage du segment ST sur le moniteur sans vérification du réglage du filtre passe-haut.
- Le calcul de la VFC sur tous les intervalles RR, y compris ceux qui entourent les extrasystoles ou les artefacts.
- La prise d'une notification de montre connectée pour un diagnostic de fibrillation atriale. La suspicion issue de la PPG se confirme par un ECG.
- La lecture de la sensibilité et de la spécificité comme une preuve de bénéfice clinique. Le bénéfice se démontre par des résultats cliniques chez les patients dans des essais randomisés.

À retenir

- La fréquence d'échantillonnage doit dépasser le double de la plus haute fréquence utile. En pratique, on choisit une marge plus grande.
- Le nombre de bits du convertisseur fixe la plus petite différence de tension qui peut être distinguée.
- Les filtres nettoient le signal, mais ils peuvent aussi le déformer. Leurs réglages font partie de l'interprétation.
- La VFC se calcule sur des intervalles NN propres. Le SDNN et le RMSSD sont le point de départ.
- Les formats ouverts (EDF, WFDB, DICOM, HL7 aECG) conservent ensemble les échantillons et les métadonnées nécessaires à l'interprétation.
- L'IA appliquée à l'ECG a augmenté le nombre de diagnostics dans des essais randomisés, mais le bénéfice pronostique et le comportement dans d'autres populations doivent être démontrés séparément.

Questions d'autoévaluation

1. (QCS) Le standard minimal IFCN-ILAE (2023) pour l'EEG de routine recommande une fréquence d'échantillonnage minimale de :

a) 70 Hz
b) 100 Hz
c) 140 Hz
d) 256 Hz
e) 1 000 Hz

2. (QCS) Une interférence sinusoïdale de 50 Hz est échantillonnée à 60 Hz, sans filtre anti-repliement. Dans le signal numérique, elle apparaîtra comme une composante de :

a) 50 Hz
b) 110 Hz
c) 10 Hz
d) 60 Hz
e) 0 Hz

3. (QCS) Dans l'essai randomisé EAGLE, l'accès des médecins de famille au résultat d'un algorithme d'IA appliqué à l'ECG :

a) a réduit la mortalité à un an
b) a augmenté la proportion de nouveaux diagnostics de fraction d'éjection basse (2,1 % contre 1,6 %)
c) a réduit de moitié le nombre total d'échocardiographies
d) a remplacé l'échocardiographie comme méthode diagnostique
e) n'a pas modifié le nombre de nouveaux diagnostics

4. (QCM) Concernant la variabilité de la fréquence cardiaque, les affirmations correctes sont :

a) le RMSSD se calcule à partir des différences entre intervalles RR successifs
b) le SDNN est l'écart-type des intervalles entre battements sinusaux normaux
c) les battements ectopiques sont conservés sans modification dans la série utilisée pour le calcul
d) la bande HF est définie typiquement entre 0,12 et 0,40 Hz et reflète l'influence de la respiration
e) la VFC peut être calculée à partir des valeurs de SpO2 affichées par l'oxymètre de pouls

5. (QCM) Concernant la détection de la fibrillation atriale par les montres connectées, les affirmations correctes sont :

a) une suspicion fondée sur la PPG doit être confirmée par un ECG
b) dans la Fitbit Heart Study, la valeur prédictive positive d'une notification répétée pendant la surveillance ECG était de 98,2 %
c) le dépistage continu par dispositifs portables a nettement réduit les événements indésirables chez les personnes asymptomatiques
d) une méta-analyse de 2025 a trouvé une hétérogénéité très importante entre les études d'exactitude
e) la probabilité de confirmation de la fibrillation atriale ne dépend pas de la durée de la surveillance ECG

CHAPITRE 7

Chapitre 7. Imagerie médicale numérique : DICOM, PACS et analyse assistée

Objectifs

À la fin du chapitre, l'étudiant sera capable :

- de calculer la taille non compressée d'une série TDM à partir de la matrice, de la profondeur de bits et du nombre d'images ;
- de convertir une valeur exprimée en unités Hounsfield en niveau de gris affiché, pour une fenêtre définie par son niveau et sa largeur ;
- de décrire le modèle d'information DICOM (patient, étude, série, instance) et de reconnaître les principaux attributs de l'en-tête d'un fichier ;
- d'expliquer le parcours d'un examen à travers le RIS, le PACS et le dossier électronique, ainsi que le rôle du profil IHE Scheduled Workflow ;
- de comparer la compression sans perte et la compression avec perte du point de vue de la sécurité diagnostique ;
- d'interpréter de manière critique un résultat publié sur l'intelligence artificielle en imagerie, en tenant compte de la validation externe et du biais d'automatisation.

7.1 L'image numérique

7.1.1 Pixel, voxel, matrice

Pour l'ordinateur, une image médicale est un tableau de nombres. Chaque élément du tableau bidimensionnel s'appelle pixel (*picture element*) et représente l'intensité d'une petite surface carrée de l'image. Le pixel correspond à un élément de volume du corps examiné, appelé voxel (*volume element*), et un volume de tomodensitométrie (TDM) ou d'imagerie par résonance magnétique (IRM) se représente comme un tableau tridimensionnel, c'est-à-dire comme une pile de coupes [24]. Les dimensions du voxel comptent. Si l'épaisseur de coupe diffère du côté du pixel dans le plan, les voxels sont anisotropes, et les reconstructions coronales ou sagittales obtenues à partir des images axiales perdent en netteté. La plupart des appareils de TDM et d'IRM modernes produisent toutefois des voxels isotropes [24].

Le nombre de lignes et de colonnes forme la matrice de l'image. Elle varie beaucoup d'une modalité à l'autre : de 64 x 64 pixels pour certaines images de médecine nucléaire jusqu'à environ 5 000 x 4 000 pixels pour la mammographie [68].

7.1.2 Profondeur de bits et résolution

Chaque pixel stocke un nombre codé sur un certain nombre de bits. Avec 8 bits, on peut représenter 2^8 = 256 valeurs, avec 12 bits, 4 096, et avec 16 bits, 65 536. Pour les images TDM, le standard DICOM impose 16 bits alloués par pixel, dont 12 à 16 bits contiennent effectivement la valeur [245]. Les images en couleur (lames de pathologie numérisées, photographies

dermatologiques) utilisent en règle générale trois canaux, rouge, vert et bleu, chacun avec sa propre profondeur.

La qualité d'une image numérique se décrit par trois grandeurs. La résolution spatiale indique dans quelle mesure deux points proches sont séparés et dépend du nombre de pixels par unité de surface. La résolution en contraste indique la plus petite différence d'intensité que l'on distingue encore, et le nombre de bits par pixel la limite. La résolution temporelle, enfin, représente le temps nécessaire pour former une image : à partir d'environ 30 images par seconde, le cœur en mouvement apparaît sans flou [24]. Les microcalcifications d'une mammographie exigent une résolution spatiale élevée. Une lésion hépatique isodense exige une bonne résolution en contraste.

7.1.3 Quelle place occupe une série TDM : exemple de calcul

Le calcul se fait en trois étapes et mérite qu'on s'y exerce, car il explique pourquoi les hôpitaux ont besoin d'archives dédiées.

1. Une image TDM axiale a une matrice de 512 x 512 = 262 144 pixels.
2. Avec 16 bits (2 octets) par pixel, une image occupe 262 144 x 2 = 524 288 octets, soit environ 0,52 Mo, sans en-tête et sans compression.
3. Une série axiale thorax-abdomen-pelvis reconstruite à 1 mm, sur une longueur de 60 cm, compte environ 600 images : 600 x 0,52 Mo ≈ 315 Mo.

Le même examen est souvent reconstruit plusieurs fois (filtre pour les tissus mous, filtre pour le poumon, plans coronaux), et le volume par patient se multiplie. Les valeurs indicatives publiées sont de l'ordre de 500 images et 250 Mo pour un examen TDM, 63 Mo pour une IRM et 153 Mo pour une mammographie, étant entendu que le nombre d'images et la résolution augmentent d'année en année [68].

7.1.4 L'échelle de Hounsfield et la fenêtre de visualisation

Les valeurs numériques des pixels TDM s'expriment en unités Hounsfield (UH ; *Hounsfield units*, HU). L'échelle est définie de telle sorte que l'eau vaut 0 UH et l'air environ -1 000 UH. En pratique, les valeurs s'étendent de -1 000 à environ +3 000 [24]. Dans le fichier DICOM, la valeur stockée (SV) se convertit en UH par la relation UH = m · SV + b, où m est le Rescale Slope (0028,1053) et b le Rescale Intercept (0028,1052) [245]. Par exemple, avec m = 1 et b = -1 024, une valeur stockée de 1 064 correspond à 40 UH.

L'œil humain ne distingue pas plus d'une centaine de nuances de gris. Un moniteur en affiche habituellement 256. Si les quelque 4 000 valeurs UH étaient toutes comprimées linéairement sur 256 niveaux, chaque niveau de gris couvrirait environ 16 UH, et les petites différences de densité entre deux tissus mous deviendraient difficiles à voir. C'est pourquoi on utilise la fenêtre de visualisation, ou fenêtrage (*windowing*) : le médecin choisit le niveau L (*window level* ou *center*, centre de l'intervalle) et la largeur W (*window width*), et seules les valeurs de l'intervalle [L - W/2 ; L + W/2] sont étalées sur toute l'échelle de gris [24]. Les valeurs situées sous l'intervalle apparaissent noires, celles qui se trouvent au-dessus, blanches. Le standard DICOM stocke la fenêtre recommandée dans les attributs Window Center (0028,1050) et Window Width (0028,1051), avec une fonction linéaire par défaut [245].

La forme simplifiée de la transformation est : G = 255 · (UH - (L - W/2)) / W, G étant limité à l'intervalle 0-255.

Exemple numérique. La fenêtre médiastinale, avec L = 40 et W = 400, couvre l'intervalle de -160 à +240 UH. Un ganglion de 20 UH reçoit 255 · (20 + 160) / 400 ≈ 115, soit un gris moyen.

Le parenchyme pulmonaire, d'environ -850 UH, tombe sous -160 et apparaît noir. La fenêtre pulmonaire, avec L = -600 et W = 1 500, couvre de -1 350 à +150 UH. Le même parenchyme de -850 UH reçoit maintenant 255 · 500 / 1 500 = 85, un gris foncé, sur lequel on voit les vaisseaux et un éventuel nodule. Les tissus mous de 40 UH deviennent presque blancs (237) et ne peuvent plus être comparés entre eux. La figure 7.1 montre les deux fonctions de transfert et leur effet sur un fantôme synthétique.

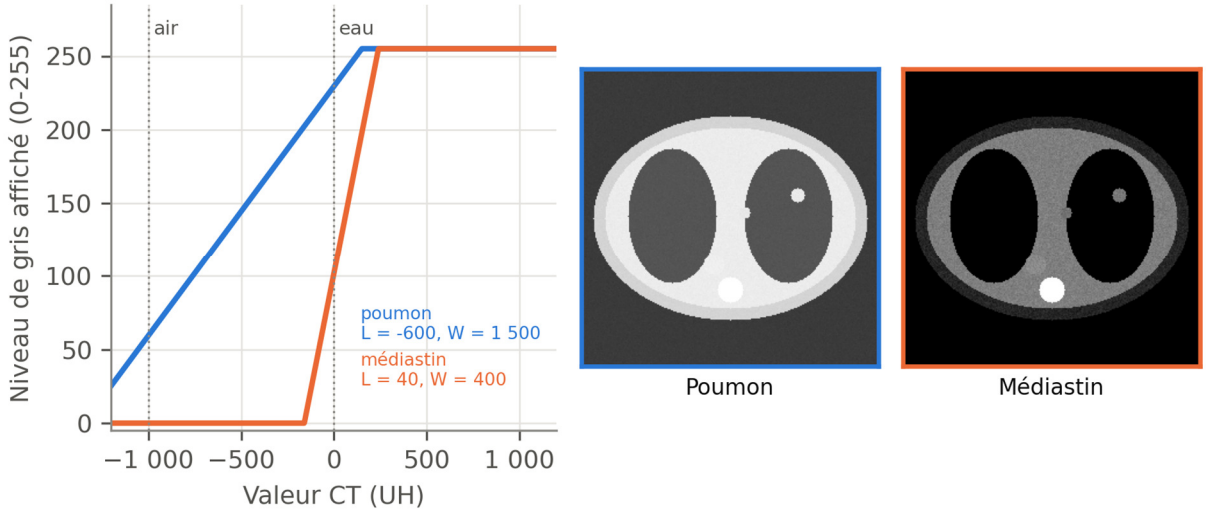


Figure 7.1. Conversion des valeurs TDM (UH) en niveaux de gris pour la fenêtre pulmonaire (L = -600, W = 1 500) et la fenêtre médiastinale (L = 40, W = 400). À droite, la même coupe synthétique affichée dans les deux fenêtres. Données synthétiques, illustratives ; les valeurs UH du fantôme sont choisies à des fins didactiques.

La fenêtre ne modifie pas les données. Elle ne change que l'affichage, et une lésion "invisible" dans une fenêtre peut être évidente dans une autre.

7.2 Les modalités d'imagerie du point de vue informatique

Du point de vue de l'informaticien, les modalités diffèrent par ce qu'elles mesurent, par la taille et le nombre des images et par le type de données associées. La radiographie projette sur un seul plan toutes les structures traversées par le faisceau, possède une résolution spatiale très élevée et une résolution en contraste plus faible que la TDM ou l'IRM. Presque tous les services sont passés du film à la radiographie numérique, soit avec des plaques au phosphore lues dans un second temps (CR), soit avec des détecteurs qui produisent directement l'image (DR) [24]. La tomodensitométrie (TDM ; *computed tomography*, CT) calcule la densité de chaque voxel au moyen d'algorithmes de reconstruction. L'image TDM est donc une image calculée, et non photographiée. L'échographie produit des images numériques dès le départ, souvent sous forme de séquences dynamiques (*cine loops*), auxquelles s'ajoute l'information Doppler sur le flux. L'imagerie par résonance magnétique (IRM) génère plusieurs séquences (T1, T2 et d'autres) pour la même région, tandis que la médecine nucléaire et la TEP rendent compte de la distribution d'un radiotraceur, avec une information fonctionnelle et une résolution spatiale réduite, raison pour laquelle elles sont fréquemment fusionnées avec la TDM [24].

La tomosynthèse mammaire numérique (*digital breast tomosynthesis*, DBT) acquiert plusieurs projections à faible dose du sein comprimé, à partir desquelles on reconstruit des coupes que l'on peut parcourir couche par couche à la station de travail [246]. La pathologie numérique numérise les lames de verre dans leur intégralité (*whole slide imaging*, WSI). Une telle image est en couleur, organisée en pyramide de "tuiles" (*tiles*) à plusieurs grossissements, et occupe entre 500

Mo et 5 Go par lame, contre 5 à 200 Mo pour une étude radiologique typique [247]. Avant l'utilisation diagnostique, chaque laboratoire valide le système WSI sur au moins 60 cas, en comparant le diagnostic posé à l'écran avec celui posé sur lame [248].

Tableau 7.1. Caractéristiques des données produites par les principales modalités. Valeurs indicatives. Source : [24,68,246,247].

Modalité	Ce qu'elle produit	Images / examen	Taille / examen
Radiographie numérique (CR/DR)	projection 2D, résolution spatiale élevée	environ 3	environ 29 Mo
TDM	volume de coupes en UH	environ 500	environ 250 Mo
IRM	plusieurs séquences (T1, T2, etc.)	environ 500	environ 63 Mo
Échographie	images et séquences ciné, Doppler	environ 50	environ 25 Mo
Mammographie 2D	4 incidences, matrice très grande	4	environ 153 Mo
Tomosynthèse (DBT)	coupes reconstruites à partir de projections multiples	coupes multiples	variable
Médecine nucléaire, TEP	distribution du radiotraceur, souvent fusionnée avec la TDM	environ 25 (scintigraphie)	environ 1 Mo (scintigraphie)
Pathologie numérique (WSI)	image couleur de l'ordre du gigapixel, pyramide de tuiles	1 par lame	500 Mo - 5 Go par lame

7.3 La compression des images

Les volumes du tableau 7.1 expliquent l'intérêt porté à la compression. Il existe deux familles de méthodes. La compression sans perte (*lossless*) code la redondance de l'image de telle sorte que l'original se reconstitue exactement, bit par bit. La compression avec perte (*lossy*, irréversible) produit une image seulement semblable visuellement à l'original, avec un degré de fidélité choisi par l'utilisateur. Les méthodes sans perte atteignent des taux de compression d'environ 2,5:1, tandis que celles avec perte peuvent aller, pour certains types d'images, jusqu'à 40:1 sans perte perceptible ou diagnostique [68]. La série TDM de 315 Mo de l'exemple ci-dessus occuperait, compressée sans perte, environ 126 Mo.

Le problème de la compression avec perte est qu'il n'existe pas de réglage universel qui garantisse à la fois un bon taux et l'absence d'altérations visibles. Les méthodes se règlent en fonction de la tâche diagnostique concernée, et la compression est plus largement acceptée en cardiologie et en pathologie qu'en radiologie. À mesure que le prix du stockage baisse plus vite que ne croît le nombre d'examens, la pression en faveur de la compression irréversible diminue [68]. Le standard DICOM permet les deux variantes au moyen des syntaxes de transfert (JPEG, JPEG-LS, JPEG 2000 et d'autres), et le fichier conserve l'information sur le type de compression appliqué [245].

Les données récentes sont nuancées. Dans une étude portant sur 9 111 mammographies, les modèles d'apprentissage profond ont conservé leurs performances en cas de compression modérée, mais l'aire sous la courbe ROC est passée de 0,87 à 0,79 pour des taux extrêmes, de 10 000-11 000:1 [249]. Douze radiologues et manipulateurs ont jugé des mammographies au format JPEG avec une compression minimale équivalentes à l'original DICOM, mais cette conclusion

concerne l'utilisation en formation, et non en diagnostic [250]. En pathologie numérique, la compression JPEG avec perte est déjà courante et peut introduire des différences de couleur et de texture, et les compresseurs fondés sur des réseaux de neurones se transfèrent mal d'un schéma de compression à l'autre [251]. Règle pratique pour l'étudiant : l'archive diagnostique conserve l'image originale ou compressée sans perte, et la compression irréversible n'est utilisée que là où un protocole validé l'autorise.

7.4 Le standard DICOM

7.4.1 Origine et statut

Les premiers appareils de TDM et d'IRM enregistraient les données sur bandes magnétiques, dans des formats propres à chaque fabricant, que les chercheurs devaient déchiffrer eux-mêmes [68]. Le Collège américain de radiologie (ACR) et l'Association nationale des fabricants d'équipements électriques (NEMA) ont publié en 1985 le standard ACR-NEMA, remplacé en 1993 par DICOM (*Digital Imaging and Communications in Medicine*), qui a apporté la communication en réseau TCP/IP, l'échange sur supports physiques (CD, DVD, USB) et une description formelle des services [252]. Les parties du standard sont numérotées de PS3.1 (introduction) à PS3.22 (communication en temps réel), deux d'entre elles (PS3.9 et PS3.13) ayant été retirées. Parmi les autres, PS3.3 décrit les objets d'information, PS3.4 les services, PS3.6 le dictionnaire de données et PS3.18 les services web [252]. Les suppléments approuvés sont consolidés plusieurs fois par an dans de nouvelles éditions. Au 27 septembre 2026, l'édition courante publiée sur le site officiel est la 2026d [245].

DICOM est également repris comme norme internationale ISO. Le projet ISO/DIS 12052:2025, soumis au vote de mars à juin 2025, annonçait la troisième édition, qui remplace l'ISO 12052:2017 [252]. L'édition publiée est l'ISO 12052:2026 [253]. Tout produit qui déclare sa conformité à DICOM doit disposer d'une déclaration de conformité (*conformance statement*) qui précise exactement quels services et quels objets il prend en charge [252]. Lors de l'achat d'un appareil, l'hôpital compare ces déclarations.

7.4.2 Le modèle d'information

La plupart des objets DICOM suivent un modèle commun comprenant les entités patient, étude, série, équipement, référentiel spatial et l'instance proprement dite (figure 7.2). Une étude correspond approximativement à une procédure demandée, et une série, à une étape d'acquisition effectuée [252]. Chaque instance, par exemple une image, reçoit un identifiant unique au niveau mondial (UID), qui l'accompagne dans tous les échanges. Si le sens du contenu est modifié, on crée une nouvelle instance, avec un nouvel UID [252].

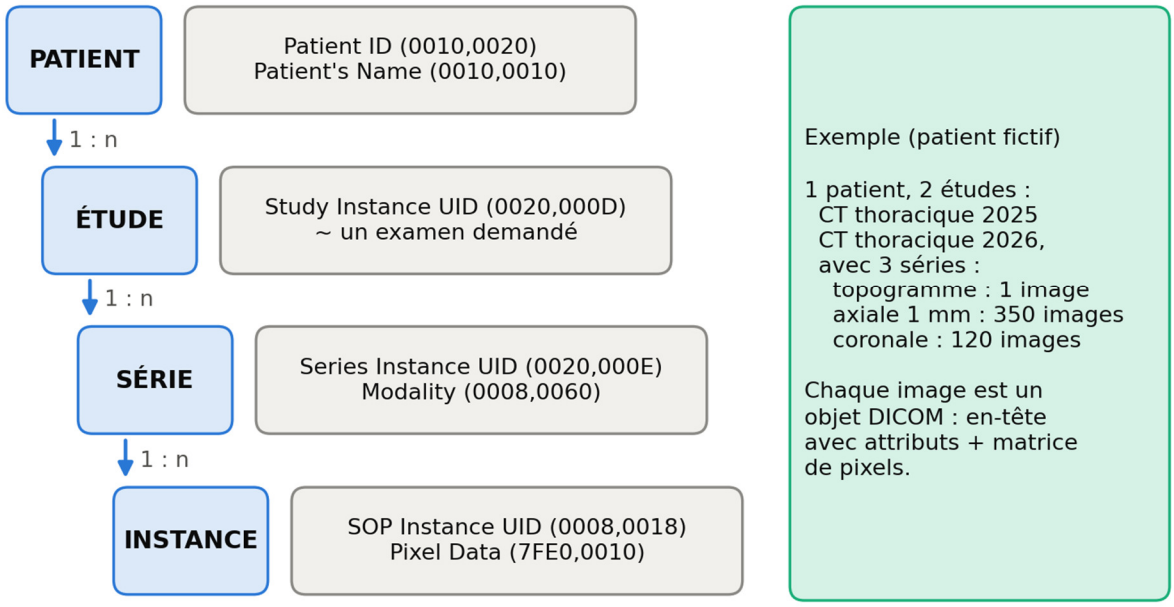


Figure 7.2. Hiérarchie du modèle d'information DICOM : un patient a une ou plusieurs études, une étude a une ou plusieurs séries, une série contient des instances (images). À droite, un exemple fictif. Les attributs et les étiquettes sont repris de PS3.6 [245].

7.4.3 Attributs, étiquettes, IOD et SOP

Un fichier DICOM comporte un en-tête formé d'attributs et, généralement, une matrice de pixels. Chaque attribut possède une étiquette (*tag*) formée de deux nombres hexadécimaux, le groupe et l'élément, un nom et un type de valeur (*value representation*, VR). Voici quelques exemples tirés du dictionnaire de données : (0010,0010) Patient's Name, de type PN (nom de personne), (0010,0020) Patient ID, (0008,0060) Modality, (0020,000D) Study Instance UID, (0028,0030) Pixel Spacing, qui donne la taille du pixel en millimètres, et (7FE0,0010) Pixel Data, qui contient l'image elle-même [245]. Pixel Spacing rend possibles les mesures en millimètres, ce que vous vérifierez au TP 7.

La définition d'un objet d'information (*Information Object Definition*, IOD) décrit de manière abstraite un type d'objet réel, par exemple "image TDM" ou "rapport structuré", avec ses attributs. La combinaison d'un IOD et d'un service forme une classe SOP (*Service-Object Pair*), par exemple "stockage d'une image TDM" [245,252]. La classe SOP d'une instance est inscrite dans l'attribut SOP Class UID (0008,0016) [245].

7.4.4 Les services réseau et DICOMweb

Les services DICOM classiques fonctionnent sur TCP/IP, entre deux applications qui établissent une association. Le tableau 7.2 résume les services qu'un étudiant rencontre le plus souvent.

Tableau 7.2. Services DICOM courants et leurs équivalents DICOMweb. Source : [245,254,255].

Service	Ce qu'il fait	Exemple d'utilisation	Équivalent web
C-ECHO	vérifie la liaison entre deux nœuds	test lors de l'installation d'un appareil	-
C-STORE	envoie et stocke une instance	l'appareil TDM envoie la série au PACS	STOW-RS
C-FIND	recherche selon des critères	toutes les études du patient X	QIDO-RS
C-MOVE / C-GET	récupère les instances trouvées	la station récupère l'étude antérieure	WADO-RS
Modality Worklist	liste de travail pour l'appareil	données du patient reprises du RIS	UPS-RS
MPPS	signale l'état de la procédure	"examen terminé" vers le RIS	UPS-RS

DICOMweb réunit les services RESTful définis dans PS3.18 : QIDO-RS pour la recherche, WADO-RS pour la récupération, STOW-RS pour le stockage et UPS-RS pour les listes de travail, avec une représentation des attributs en JSON ou en XML [254]. L'avantage est qu'un visualiseur web ou une application mobile peut accéder aux images avec les outils habituels des développeurs web. Un flux construit uniquement sur des interfaces FHIR et DICOMweb a été décrit pour la demande d'examen, la programmation, l'acquisition, la visualisation et le compte rendu [256]. En pathologie numérique, WADO-RS permet de demander individuellement les tuiles d'une image de plusieurs gigaoctets, sans télécharger le fichier entier [247]. Lors du Connectathon DICOM WSI de 2025, 32 implémentations (scanners, archives, visualiseurs, sources d'annotations) ont échangé des images et des annotations au moyen des mécanismes standard de DICOM et de DICOMweb [257].

7.4.5 L'anonymisation

L'en-tête DICOM contient des données personnelles, et parfois les pixels en contiennent aussi : le nom du patient "incrusté" dans l'image d'une échographie ou d'une capture d'écran, ou encore les traits du visage reconstructibles à partir d'une TDM crânienne. La partie PS3.15 définit le profil de base de confidentialité (*Basic Application Level Confidentiality Profile*), qui supprime ou remplace les attributs identifiants, avec des options supplémentaires telles que le nettoyage des données de pixels (*Clean Pixel Data*), le nettoyage des traits visuels reconnaissables et la conservation des relations temporelles longitudinales [245]. Cette dernière option compte en recherche, où l'intervalle entre deux examens doit être conservé même si les dates réelles ne le sont pas.

Un atelier organisé en 2023 par l'Institut national du cancer des États-Unis a réuni les recommandations du groupe de travail MIDI (*Medical Image De-Identification Initiative*), les outils disponibles et les approches internationales en matière d'anonymisation des images destinées à la recherche et au partage, y compris pour les applications d'IA [258]. L'anonymisation fait partie d'un ensemble plus large de mesures de sécurité pour le PACS : transport chiffré, chiffrement sélectif de l'en-tête, signatures numériques et absence d'exposition directe des serveurs PACS sur Internet [259]. Les aspects juridiques (RGPD, pseudonymisation) sont traités dans le chapitre consacré à la protection des données.

7.5 PACS, RIS, VNA et flux de travail radiologique

7.5.1 Les composants

Le système d'archivage et de transmission d'images (*Picture Archiving and Communication System*, PACS) reçoit, stocke, distribue et affiche les images. Le système d'information radiologique (*Radiology Information System*, RIS) gère la partie textuelle : la programmation, l'état de chaque examen et le compte rendu. Là où il existe un système d'information hospitalier ou un dossier de santé électronique (DSE ; *electronic health record*, EHR), la demande d'examen et la facturation passent habituellement par celui-ci, et le RIS reprend ces fonctions surtout dans les centres d'imagerie ambulatoires [68]. Le compte rendu est dicté, et la reconnaissance vocale a réduit le délai d'obtention du compte rendu final d'une semaine, durée fréquente auparavant, à moins d'une heure [68].

Les premiers PACS liaient étroitement l'archive aux stations d'affichage du même fabricant. Lors d'un changement de système, la migration des données exigeait la coopération de l'ancien fournisseur. L'archive neutre vis-à-vis des fabricants (*vendor neutral archive*, VNA) sépare le

stockage de la visualisation, utilise DICOM pour l'accès et peut aussi héberger des images extérieures à la radiologie. Dans les grands hôpitaux, le VNA est devenu la règle [68]. Il existe aussi des PACS à code source ouvert. Une évaluation comparative a désigné Orthanc, DCM4CHE, DCMTK et Dicoogle comme les plus matures [260].

7.5.2 Le flux de travail et le profil IHE Scheduled Workflow

La figure 7.3 suit un examen TDM de la demande jusqu'au résultat. Le médecin traitant émet la demande dans le DSE. Le RIS programme l'examen et le publie dans la liste de travail de l'appareil (*Modality Worklist*), de sorte que le manipulateur n'a plus à saisir le nom du patient et que les données démographiques arrivent correctement dans l'en-tête DICOM. L'appareil signale le début et la fin de la procédure (*Modality Performed Procedure Step*, MPPS), envoie les images au PACS et demande la confirmation de leur prise en charge (*storage commitment*). Le radiologue ouvre l'étude à la station de lecture, la compare aux examens antérieurs et signe le compte rendu, qui parvient au RIS et au DSE, généralement avec un lien vers les images.

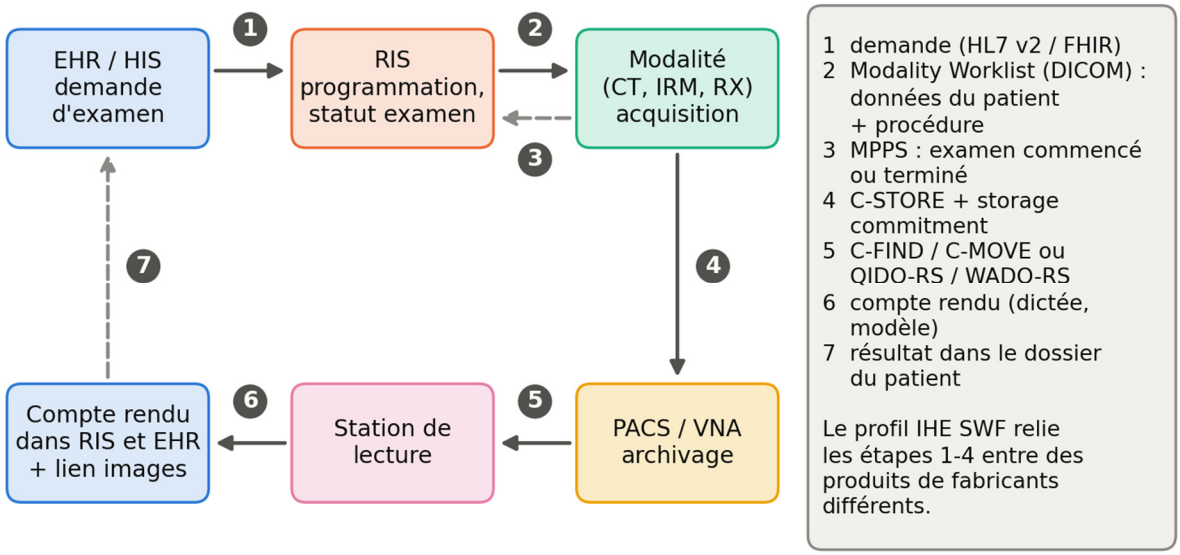


Figure 7.3. Flux d'un examen dans le réseau DSE-RIS-modalité-PACS, avec les messages HL7/FHIR et les services DICOM pour chaque étape. Adapté de [68,255].

Les standards seuls ne garantissent pas la coopération entre produits différents. L'initiative IHE (*Integrating the Healthcare Enterprise*) définit des profils d'intégration qui précisent quels standards et quelles options utiliser dans un scénario donné. Le profil Scheduled Workflow (SWF) intègre la demande, la programmation, l'acquisition, le stockage et la visualisation des examens radiologiques. Il utilise la liste de travail, le MPPS et le storage commitment, ce dernier service transférant la responsabilité des images de l'appareil au PACS et prévenant leur perte [255]. Les profils n'ont pas le droit de contredire le standard DICOM ; ils peuvent seulement ajouter des exigences [252]. La révision 23.0 du cadre technique IHE pour la radiologie, publiée le 8 août 2025, inclut la variante SWF.b et des profils en cours de test pour les résultats des algorithmes d'IA [261].

7.5.3 La téléradiologie et le PACS dans le cloud

La téléradiologie désigne l'interprétation des images en un lieu différent de celui de l'acquisition. Elle a commencé par la transmission nocturne des examens au radiologue de garde

resté à domicile et elle est devenue une activité organisée, avec des équipes qui interprètent pour plusieurs hôpitaux. Les obstacles principaux ne sont plus techniques ; ils tiennent à l'autorisation d'exercice et à l'accréditation du médecin dans la juridiction du patient [68]. Les aspects organisationnels et juridiques de la télémédecine sont discutés au chapitre 9.

Les économies d'échelle poussent de nombreux petits prestataires de services d'imagerie vers le stockage dans le cloud, et les algorithmes d'analyse gourmands en calcul ont poussé dans la même direction [68]. Une architecture PACS native du cloud, neutre vis-à-vis des fabricants, dotée d'un dispositif local qui conserve l'interface DICOM classique, stockait en production plus de 450 To d'images et desservait plus de 4 000 nœuds DICOM [262]. Le passage au cloud ne dispense pas l'hôpital de sa responsabilité en matière de confidentialité, de contrats et de continuité de l'accès.

7.6 Traitement et analyse des images

7.6.1 Opérations globales : histogramme, contraste, filtres

Une chaîne de traitement typique comprend les opérations globales, l'amélioration de l'image, la visualisation, la quantification, la segmentation, le recalage (alignement de deux images) et, pour finir, la classification [24]. L'histogramme indique le nombre de pixels pour chaque valeur d'intensité, et c'est la première chose à regarder sur une nouvelle image. L'égalisation d'histogramme redistribue les valeurs les plus fréquentes afin d'augmenter le contraste global. Le filtrage par convolution applique une petite matrice de coefficients (noyau) à chaque pixel et à ses voisins. Un filtre gaussien lisse l'image et réduit le bruit, tandis qu'un filtre médian élimine les points isolés en préservant mieux les contours. L'accentuation par masque flou (*unsharp masking*) augmente le contraste local et accentue les détails fins [24]. La fenêtre TDM de la section 7.1 est elle aussi une opération globale.

7.6.2 Segmentation et mesures

La segmentation délimite une structure : un organe, une tumeur, un nodule. La méthode la plus simple est le seuillage, par lequel les pixels situés au-dessus d'une valeur deviennent "objet" et les autres "fond". Viennent ensuite les méthodes fondées sur les régions, sur les contours et sur le regroupement (*clustering*) et, ces dernières années, les réseaux de neurones profonds [24]. nnU-Net est une méthode qui configure d'elle-même le prétraitement, l'architecture et l'entraînement pour chaque nouveau jeu de données ; elle a surpassé la plupart des solutions spécialisées sur 23 jeux de données publics issus de compétitions internationales [263]. Construit sur nnU-Net, TotalSegmentator segmente 104 structures anatomiques en TDM, avec un coefficient de Dice de 0,943 sur le jeu de test [264]. Le coefficient de Dice mesure le chevauchement entre la segmentation automatique et la segmentation de référence, de 0 (aucun chevauchement) à 1 (identité).

Après la segmentation viennent les mesures : longueurs, aires, volumes, densité moyenne en UH. Toutes dépendent de la calibration spatiale inscrite dans l'en-tête (Pixel Spacing, épaisseur de coupe). Un fichier exporté en JPEG perd cette information et, avec elle, la possibilité de mesurer en millimètres.

7.6.3 La radiomique en bref

La radiomique extrait des images un grand nombre de caractéristiques quantitatives (forme, statistiques d'intensité, texture) et les analyse par des méthodes statistiques ou d'apprentissage

automatique, par exemple pour la classification de la maladie ou la stratification du risque [265]. Le flux typique comprend la segmentation de la région d'intérêt, le traitement des images (harmonisation de la taille des pixels et des niveaux de gris), l'extraction des caractéristiques et leur sélection, après quoi on construit le modèle [265]. Le point faible est la reproductibilité : les caractéristiques varient avec l'appareil, le protocole d'acquisition, la reconstruction et le logiciel utilisé [265]. L'initiative IBSI (*Image Biomarker Standardisation Initiative*), qui réunit 25 équipes, a standardisé 169 des 174 caractéristiques radiomiques, ce qui permet d'étalonner les logiciels les uns par rapport aux autres [266].

7.7 L'intelligence artificielle en imagerie, 2026

7.7.1 Combien de dispositifs, dans quelles spécialités

L'agence américaine des produits alimentaires et des médicaments (FDA) publie une liste des dispositifs médicaux qui utilisent l'IA et ont reçu une autorisation de mise sur le marché aux États-Unis. La liste n'est pas exhaustive : les dispositifs sont identifiés principalement d'après les termes liés à l'IA figurant dans les documents d'autorisation, et elle est mise à jour périodiquement [267]. La page consultée le 27 septembre 2026 avait été mise à jour le 4 septembre 2026. Le fichier téléchargeable contenait 1 614 dispositifs, avec des décisions finales comprises entre septembre 1995 et le 29 juin 2026 [267]. La radiologie était le panel principal pour 1 230 d'entre eux (76,2 %), suivie du domaine cardiovasculaire, avec 154 (9,5 %), et de la neurologie, avec 73 (4,5 %). Pour la seule année 2025, 335 autorisations ont été enregistrées (figure 7.4).

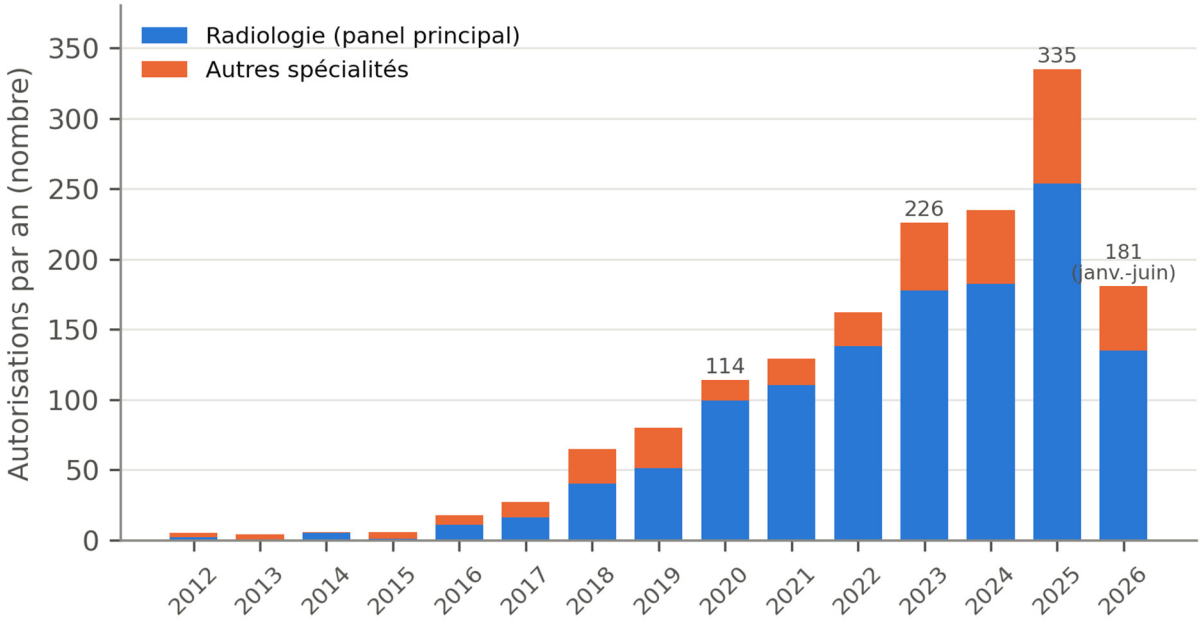


Figure 7.4. Dispositifs médicaux à intelligence artificielle autorisés par la FDA, par année de décision finale (2012-2026), radiologie par rapport aux autres spécialités (panel principal d'évaluation). Pour 2026, les décisions de janvier à juin sont incluses. Calcul propre sur la liste officielle téléchargée le 27 septembre 2026 (page mise à jour le 4 septembre 2026) [267].

Une analyse longitudinale des 1 430 dispositifs autorisés jusqu'à la fin de 2025 a montré que la part de la radiologie avait atteint un maximum de 85,5 % en 2020-2022 et était redescendue à 77,5 % en 2023-2025, première baisse statistiquement significative du registre [268]. La liste actuelle donne, pour la même période, 1 433 dispositifs, la faible différence provenant des mises à jour ultérieures. Un grand nombre d'autorisations ne signifie pas pour autant des preuves cliniques

solides. Dans la même analyse, 96,2 % des dispositifs avaient été autorisés par la procédure 510(k), qui repose sur l'équivalence avec un dispositif déjà présent sur le marché [268].

7.7.2 Le dépistage mammographique : preuves randomisées et données de la pratique

Pour la mammographie de dépistage, il existe déjà des preuves issues d'essais randomisés. L'essai suédois MASAI a randomisé des femmes du programme national entre une lecture assistée par l'IA et la double lecture standard. Le système d'IA attribuait à chaque examen un score de risque de 1 à 10. Les examens de score 1-9 étaient confiés à un seul radiologue, ceux de score 10 à deux radiologues. Les lecteurs voyaient le score et, pour les scores 8-10, également les marques de l'algorithme sur l'image. L'analyse de sécurité portant sur 80 033 femmes a trouvé un taux de détection de 6,1 pour 1 000 avec l'IA contre 5,1 pour 1 000 sans IA, avec une réduction de 44,3 % de la charge de lecture [269]. Sur l'ensemble de la cohorte de 105 934 femmes, la détection a été de 6,4 contre 5,0 pour 1 000 (rapport 1,29), sans augmentation significative des rappels ni des résultats faux positifs, et les cancers supplémentaires étaient surtout des cancers invasifs de petite taille, sans atteinte ganglionnaire [270]. Le critère de jugement principal, le taux de cancers de l'intervalle (survenus entre deux campagnes de dépistage), a été publié ultérieurement : 1,55 contre 1,76 pour 1 000, soit une non-infériorité, avec une sensibilité de 80,5 % contre 73,8 % et la même spécificité, de 98,5 %, dans les deux bras [271].

L'étude allemande PRAIM a analysé la mise en œuvre en pratique réelle, dans 12 centres, auprès de 463 094 femmes et de 119 radiologues qui ont choisi eux-mêmes d'utiliser ou non le système d'IA. La détection a été de 6,7 pour 1 000 avec l'IA contre 5,7 pour 1 000 sans, et le taux de rappel n'a pas augmenté (37,4 contre 38,3 pour 1 000) [272]. PRAIM est une étude observationnelle ; les groupes n'y sont donc pas aussi strictement comparables que dans un essai contrôlé randomisé (ECR).

7.7.3 Les nodules pulmonaires en TDM

Pour les programmes de dépistage pulmonaire par TDM, une revue systématique des logiciels d'IA porteurs du marquage CE a inclus 11 études portant sur 19 770 patients, toutes à risque élevé de biais. La lecture assistée a été plus courte et a en général augmenté la sensibilité (de 5 à 20 points de pourcentage pour les nodules nécessitant une prise en charge), au prix d'une spécificité plus faible, de 3 à 7 points. Les auteurs ont estimé, pour une prévalence du cancer de 0,5 %, que 150 à 750 cancers supplémentaires seraient détectés pour un million de personnes examinées, mais aussi que 59 700 à 79 600 personnes sans cancer seraient orientées inutilement vers une surveillance par TDM [273]. L'évaluation technologique étendue, portant sur 27 études, a confirmé le risque élevé de biais dans toutes les études incluses [274]. L'exemple montre pourquoi une "sensibilité plus élevée" ne signifie pas automatiquement un bénéfice pour la population.

7.7.4 Limites : généralisation, biais de spectre, confiance excessive

Un algorithme performant sur les données de l'hôpital qui l'a développé peut être médiocre dans un autre hôpital. Une revue systématique de 86 algorithmes d'apprentissage profond validés en externe a trouvé une baisse des performances pour 81 % d'entre eux, baisse substantielle (d'au moins 0,10 sur l'échelle 0-1) pour 24 % [275]. Une autre revue, portant sur 535 études en radiologie, a constaté que 98 % d'entre elles étaient rétrospectives et que, dans les 77 études qui permettaient la comparaison, les performances baissaient en moyenne de 6 % lors de la validation externe [276]. La pathologie numérique montre la même tendance. Une méta-analyse de 48 études sur lames numérisées a rapporté une sensibilité moyenne de 96,3 % et une spécificité de 93,3 %,

mais 99 % des études incluses dans la revue présentaient au moins un domaine à risque de biais élevé ou incertain, ou des problèmes d'applicabilité [277], et les différences entre institutions (coloration, scanners) restent l'un des principaux problèmes du domaine [278].

Le biais de spectre apparaît lorsque les patients de l'ensemble de test ne ressemblent pas à ceux de la pratique, par exemple lorsque les cas positifs sont des tumeurs volumineuses et évidentes, et les cas négatifs, des volontaires sains. Les performances rapportées sont alors surestimées. Les biais peuvent s'introduire à chaque étape, depuis la collecte des données jusqu'à la manière dont on choisit les métriques et dont on publie les résultats [279]. Au sens statistique, le biais désigne une erreur systématique : les résultats du modèle ne représentent pas les conditions réelles [280].

Le troisième risque concerne l'humain. Le biais d'automatisation est la tendance à accepter la suggestion d'un système automatisé même lorsqu'elle contredit d'autres éléments de preuve [280]. Dans une expérience où 27 radiologues ont lu des mammographies avec un prétendu système d'IA, la proportion d'évaluations BI-RADS correctes est passée d'environ 80 % lorsque la suggestion était correcte à 20-46 % lorsqu'elle était erronée, et les débutants se sont montrés les plus vulnérables [281]. Dans une étude sur des radiographies thoraciques, les résultats erronés de l'IA ont fait passer la proportion de faux négatifs des radiologues de 2,7 % à 33,0 %, l'effet étant atténué lorsque l'algorithme encadrait la zone suspecte [282]. Une vaste étude, avec 140 radiologues et 15 tâches sur des radiographies thoraciques, a montré que l'ancienneté et la familiarité avec l'IA ne permettent pas de prédire qui tire profit de l'assistance, alors que les erreurs de l'algorithme affectent les performances des lecteurs [283].

7.7.5 Le cadre réglementaire dans l'Union européenne

Dans l'Union européenne, un logiciel destiné à une finalité médicale, y compris un algorithme de détection sur images, est un dispositif médical et relève du règlement (UE) 2017/745 relatif aux dispositifs médicaux (MDR) [284]. Le règlement (UE) 2024/1689 sur l'intelligence artificielle (AI Act) classe comme systèmes à haut risque, entre autres, les systèmes d'IA qui sont des dispositifs médicaux soumis à une évaluation de la conformité par un organisme notifié [22], et le calendrier d'application des obligations relatives à ces systèmes a été modifié par le règlement (UE) 2026/1744 du 8 juillet 2026 (*Digital Omnibus on AI*) [197].

Pièges fréquents

- Croire que la fenêtre modifie l'image. Elle ne change que l'affichage, et les valeurs UH du fichier restent les mêmes.
- Exporter une image DICOM en JPEG pour la mesurer. On perd la calibration (Pixel Spacing), les valeurs UH et, le plus souvent, une partie de la qualité.
- Confondre l'étude et la série. Une étude TDM peut comporter plusieurs séries (topogramme, axiale, coronale), et une série compte des centaines d'images.
- Considérer un fichier comme anonymisé simplement parce que le nom a été effacé. Il reste l'identifiant, la date de naissance, le texte "incrusté" dans les pixels et le visage reconstructible à partir de la TDM.
- Lire "l'IA augmente la sensibilité" sans se demander ce qu'il advient de la spécificité et du nombre de résultats faux positifs.
- Prendre le nombre de dispositifs autorisés pour une preuve de bénéfice clinique.

À retenir

- Une image TDM de 512 x 512 pixels sur 16 bits occupe environ 0,5 Mo, et une série de 600 coupes, environ 315 Mo non compressée.
- La fenêtre (L, W) affiche l'intervalle [L - W/2 ; L + W/2] sur toute l'échelle de gris.
- DICOM organise les données sur quatre niveaux : patient, étude, série, instance, chacun avec son propre UID.
- C-STORE stocke, C-FIND recherche, C-MOVE récupère ; dans DICOMweb, STOW-RS, QIDO-RS et WADO-RS font la même chose via HTTP.
- Le RIS gère la programmation et le compte rendu, le PACS gère les images, et le profil IHE SWF les relie par la liste de travail, le MPPS et le storage commitment.
- Dans l'ECR MASAI, la lecture assistée par l'IA a détecté davantage de cancers, avec 44 % de travail de lecture en moins et un taux de cancers de l'intervalle non inférieur.
- La validation externe et le biais d'automatisation sont les deux premières questions à poser à propos de tout algorithme.

Questions d'autoévaluation

1. (QCS) Une série TDM compte 400 images de matrice 512 x 512, à 16 bits par pixel. Sa taille approximative non compressée (sans les en-têtes) est de :

a) 21 Mo
b) 52 Mo
c) 105 Mo
d) 210 Mo
e) 1,6 Go

2. (QCS) Pour une fenêtre de niveau L = 40 UH et de largeur W = 400 UH, un voxel de valeur -300 UH sera affiché :

a) en blanc
b) en gris clair
c) en gris moyen
d) en gris foncé, mais distinct du noir
e) en noir

3. (QCS) Dans le modèle d'information DICOM, l'entité qui correspond approximativement à une procédure d'imagerie demandée est :

a) le patient
b) l'étude
c) la série
d) l'instance
e) l'équipement

4. (QCM) À propos des services DICOM et DICOMweb, les affirmations correctes sont les suivantes :

a) C-STORE envoie une instance vers un autre nœud, par exemple de l'appareil TDM au PACS

b) QIDO-RS est le service web de recherche des objets DICOM

c) Modality Worklist transmet les images du PACS à la station de lecture

d) WADO-RS permet de récupérer des objets DICOM via HTTP

e) DICOMweb peut représenter les attributs en JSON

5. (QCM) Dans l'essai randomisé MASAI sur le dépistage mammographique assisté par l'IA :

a) le taux de détection du cancer a été plus élevé dans le bras avec IA

b) la charge de lecture des radiologues a diminué d'environ 44 %

c) le taux de cancers de l'intervalle a été non inférieur à celui de la double lecture standard

d) la spécificité a été beaucoup plus faible dans le bras avec IA

e) les radiologues n'ont jamais vu les marques de l'algorithme

CHAPITRE 8

Chapitre 8. Le dossier de santé électronique et l'interopérabilité

Objectifs

À la fin du chapitre, l'étudiant sera capable :

- de distinguer le dossier médical électronique, le dossier de santé électronique et le dossier personnel de santé selon qui les administre et quelles données ils contiennent ;
- d'énumérer les composants d'un dossier électronique et d'exposer, à l'aide d'un exemple, les avantages et les limites de la documentation structurée par rapport au texte libre ;
- de classer un problème d'échange de données dans l'un des quatre niveaux de l'interopérabilité et d'indiquer le standard approprié pour chaque niveau ;
- de lire un segment PID d'un message HL7 v2 et une ressource FHIR Patient au format JSON, en identifiant le nom, la date de naissance et l'identifiant du patient ;
- de décrire le calendrier de l'espace européen des données de santé pour le résumé du patient et la prescription électronique ;
- de reconnaître les schémas d'utilisation à risque du dossier électronique (copier-coller, surcharge d'alertes) et de proposer une mesure de prévention.

8.1 Le dossier électronique : concepts et effets

8.1.1 EMR, DSE (EHR), PHR

Le dossier du patient remplit plusieurs fonctions : conserver les observations, informer les collègues, former les étudiants, produire de nouvelles connaissances, permettre le suivi de l'activité et justifier les interventions [285]. Le dossier de santé électronique (DSE ; *electronic health record*, EHR) est un dépôt électronique d'informations sur l'état de santé et les soins d'une personne, doté de fonctions qui soutiennent l'acte médical et les autres utilisations légitimes des données [285]. La différence avec le dossier papier ne tient pas seulement au support. Un DSE peut vérifier les données au moment même de leur saisie (par exemple, un intervalle de doses), afficher des tendances et signaler des valeurs dangereuses.

Trois termes apparentés circulent dans la littérature. Le dossier médical électronique (*electronic medical record*, EMR) désigne habituellement la documentation d'une seule institution, cabinet ou hôpital. Le DSE rassemble des données provenant de plusieurs lieux de soins, tout au long de la vie. Les auteurs du manuel de référence du domaine observent toutefois que ces termes sont souvent employés de manière interchangeable, et que beaucoup sont passés de "medical" à "health" parce que les systèmes sont utilisés tout au long du parcours de soins, par de nombreuses catégories de professionnels [285]. Le dossier personnel de santé (*personal health record*, PHR) est contrôlé par le patient : celui-ci consulte ses résultats, reçoit des messages de l'équipe soignante et peut ajouter des données mesurées à domicile, comme la pression artérielle

ou la glycémie [285]. Les portails patients, liés au DSE d'une institution, sont apparentés au PHR et sont souvent analysés avec lui, en tant que dossiers numériques centrés sur le patient [286] (tableau 8.1).

Tableau 8.1. Dossier médical électronique, dossier de santé électronique et dossier personnel de santé. Source : [285,286].

Caractéristique	EMR (dossier médical électronique)	DSE (dossier de santé électronique)	PHR (dossier personnel)
Qui le gère	une institution (cabinet, hôpital)	le réseau de prestataires, la région ou l'État	le patient
Source des données	les cliniciens de l'institution	plusieurs institutions, par échange de données	DSE, patient, dispositifs à domicile
Finalité principale	les soins courants et la facturation	la continuité des soins, les utilisations secondaires	l'information et l'implication du patient
Exemple	le dossier du logiciel du cabinet de médecine de famille	le dossier régional ou national	le portail patient, l'application mobile

8.1.2 Composants et mode de documentation

Un DSE mature n'est pas une simple archive. Le manuel de Shortliffe décrit six composants fonctionnels : la saisie, l'agrégation et la revue des données du patient, la prescription informatisée (*computerized provider order entry*, CPOE), l'aide à la décision clinique, l'accès aux ressources de connaissances, la communication au sein de l'équipe et avec le patient, le codage et la facturation [285]. Les données proprement dites se regroupent en blocs faciles à reconnaître :

- les données démographiques et administratives (identifiant, nom, date de naissance, sexe, adresse, assurance) ;
- la liste des problèmes (diagnostics actifs et résolus, de préférence codés) ;
- les médicaments en cours et l'historique des prescriptions ;
- les allergies et les intolérances, avec le type de réaction ;
- les résultats de laboratoire et les signes vitaux ;
- les notes cliniques (anamnèse, examen, évolution, lettre de sortie) ;
- les images et les comptes rendus d'imagerie, liés au PACS décrit au chapitre 7.

La liste des problèmes est souvent incomplète. Un essai randomisé mené dans 4 systèmes de santé, avec 3 DSE différents, a utilisé une alerte qui proposait d'ajouter les problèmes manquants. Le taux d'acceptation a été de 22,1 %, et, dans le bras d'intervention, on a ajouté 4,6 fois plus de problèmes que dans le groupe témoin, sans modification des indicateurs de qualité des soins [287].

La documentation se fait de deux manières, aux avantages opposés. Le texte libre (narratif) s'écrit facilement et conserve le raisonnement clinique et les nuances, mais il se prête mal à la recherche d'informations, et les abréviations comme les variantes d'expression le rendent difficile à traiter automatiquement [285]. La documentation structurée utilise des formulaires comportant des champs et des vocabulaires contrôlés. L'ordinateur peut vérifier immédiatement la cohérence des données et les réutiliser pour des alertes, des statistiques ou la recherche. Le prix à payer est le temps : dans une étude citée par le même manuel, la saisie d'un diagnostic codé en SNOMED CT a pris jusqu'à 20 secondes [285]. Une solution pratique fréquente est la solution mixte (semi-structurée), avec des champs structurés pour les données qui seront réutilisées (allergies, médicaments, signes vitaux) et du texte libre pour le raisonnement [285].

8.1.3 Quels effets ont été documentés

Les bénéfices ne sont pas automatiques, mais ils existent. Une méta-analyse de 2026 portant sur 74 essais randomisés en soins primaires a montré que les interventions numériques de sécurité, de l'aide à la décision intégrée au DSE jusqu'à l'audit et au retour d'information, avaient amélioré les indicateurs de processus relatifs à la sécurité médicamenteuse (OR 1,46 ; IC à 95 % 1,30-1,64) et modestement réduit les événements indésirables (OR 1,17 ; IC à 95 % 1,06-1,30, estimation fondée sur seulement 12 études). Les interventions à plusieurs composantes ont mieux fonctionné que celles à une seule composante [288]. Chez les patients atteints de maladies chroniques, des effets favorables des portails et des PHR ont été rapportés surtout pour l'utilisation des services recommandés, la connaissance de la maladie, l'implication du patient et l'observance du traitement, mais seulement 16 des 81 études étaient de qualité méthodologique élevée [286].

Le coût le plus visible est le temps passé devant l'écran. Une revue systématique de 18 études (pour moitié des études observationnelles de type *time-motion*, presque toutes les autres fondées sur les journaux du système) a estimé que les médecins travaillent dans le DSE environ 37 % de leur journée de travail, et les infirmières hospitalières environ 22 % [289]. Les journaux d'audit (*audit logs*) enregistrent chaque action de l'utilisateur et permettent de telles mesures à grande échelle.

Les chiffres tirés des journaux convergent. Sur les données d'environ 155 000 médecins américains utilisant le même produit commercial, le temps moyen passé dans le DSE était de 16 minutes et 14 secondes par consultation, dont 33 % pour la revue du dossier, 24 % pour la documentation et 17 % pour les prescriptions [290]. Une comparaison portant sur 371 systèmes de santé ayant recours au même fournisseur a montré que les cliniciens des États-Unis (médecins et autres professionnels ayant leurs propres rendez-vous) travaillaient activement dans le DSE 90,2 minutes par jour, contre 59,1 minutes hors des États-Unis, et recevaient 33,8 messages par jour, contre 12,8 [291]. Le produit étant le même, la différence tient aussi au contexte d'utilisation. La tendance est à la hausse : chez 141 médecins de soins primaires d'un centre universitaire, le temps passé dans le DSE, rapporté à 8 heures de consultations, a augmenté de 28,4 minutes (7,8 %) entre 2019 et 2023, et les messages par lesquels les patients demandent un avis médical ont augmenté de 55,5 % [292]. Les études *time-motion* ajoutent une dimension que les journaux ne voient pas. Chez 47 cliniciens observés directement, le passage d'une tâche à une autre avait lieu en moyenne 1,4 fois par minute, et près de la moitié des changements de tâche impliquaient la consultation ou la saisie de données [293].

Le lien avec l'épuisement professionnel (*burnout*) est documenté, bien que presque uniquement par des études transversales. Une méta-analyse de 37 études, avec 66 556 participants, a estimé une prévalence cumulée de l'épuisement de 40,4 % chez les professionnels utilisant un DSE. Pour ceux qui travaillaient davantage dans le DSE après les heures de service, le rapport de cotes (*odds ratio*) d'épuisement était de 2,43 [294]. Dans une enquête menée auprès de 2 067 médecins de famille, une bonne utilisabilité du système (saisie des données, recherche de l'information, alertes utiles) était associée à la satisfaction à l'égard du DSE, et la satisfaction à une moindre fréquence de l'épuisement [295]. Une solution évaluée récemment à grande échelle est celle des scribes fondés sur l'intelligence artificielle, qui transcrivent et résument la consultation. Dans une étude de cohorte portant sur 8 581 cliniciens de 5 centres universitaires américains, leur adoption a été associée à 16 minutes de documentation en moins pour 8 heures de consultations, mais le temps de travail dans le DSE après les heures de service n'a pas varié de manière statistiquement significative [296].

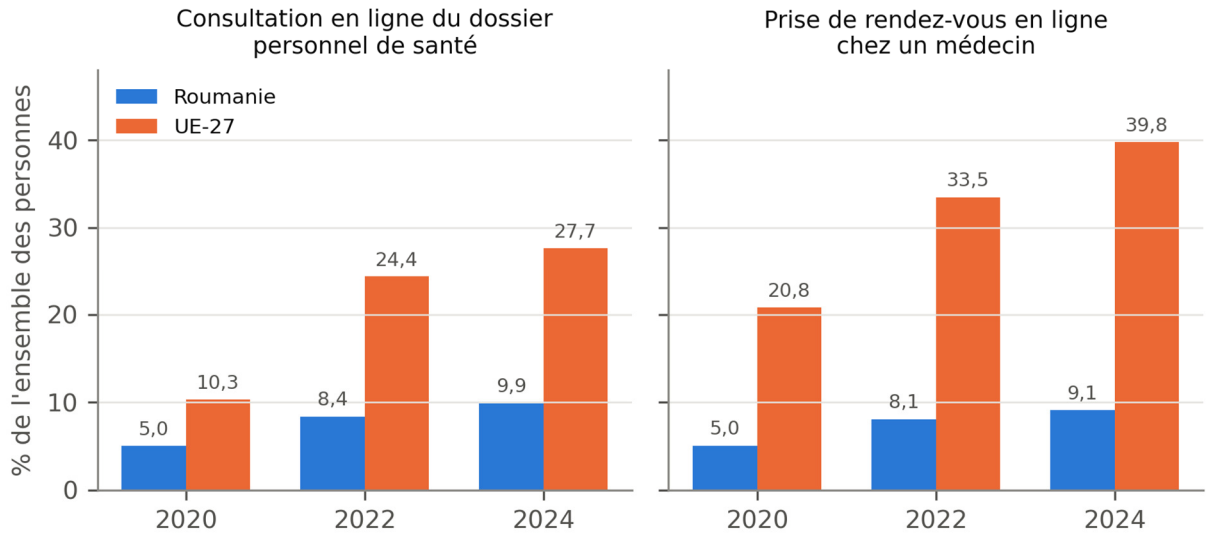


Figure 8.1. Consultation en ligne du dossier personnel de santé et prise de rendez-vous en ligne chez un médecin, Roumanie par rapport à l'UE-27, 2020-2024 (% de l'ensemble des personnes, avec la valeur de l'UE-27 pour 2020 estimée par Eurostat). Données Eurostat isoc_ci_ac_i, extraites le 27 septembre 2026 [297].

La part du patient dans l'équation reste faible en Roumanie. Selon Eurostat, en 2024, seulement 9,9 % des personnes en Roumanie avaient consulté en ligne leur dossier de santé, contre 27,7 % en moyenne dans l'UE-27, et la prise de rendez-vous en ligne chez un médecin avait été utilisée par 9,1 % d'entre elles, contre 39,8 % dans l'UE (figure 8.1) [297].

8.2 L'interopérabilité

8.2.1 Les quatre niveaux

L'interopérabilité est la capacité de deux systèmes à échanger des données et à les utiliser sans intervention manuelle. Une revue de synthèse de 24 travaux sur le DSE a retrouvé quatre niveaux : technique (fondamental), syntaxique (structurel), sémantique et organisationnel (de processus), et a désigné le niveau sémantique, c'est-à-dire la préservation du sens des données, comme le principal problème à résoudre [298]. La figure 8.2 associe à chaque niveau une question simple.

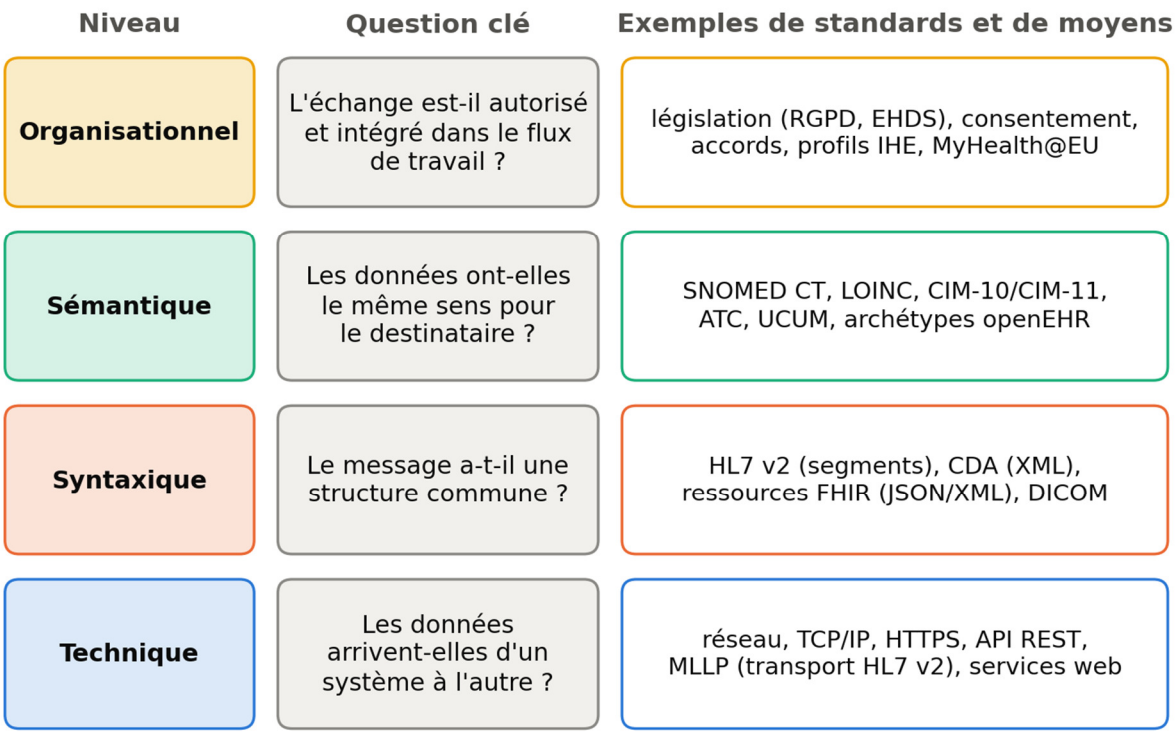


Figure 8.2. Les niveaux de l'interopérabilité, la question à laquelle répond chacun d'eux et des exemples de standards. Les niveaux supérieurs supposent que les niveaux inférieurs soient résolus. Adapté de [298].

Le niveau technique assure le transport : réseau, protocoles, interfaces de programmation. Le niveau syntaxique fixe le format, c'est-à-dire l'ordre des champs, les séparateurs et la structure du document. Le niveau sémantique exige que l'émetteur et le destinataire comprennent la même chose par le même code, ce qui suppose des terminologies communes telles que SNOMED CT pour les concepts cliniques, LOINC pour les analyses et la CIM (ICD) pour les diagnostics. Le niveau organisationnel couvre la législation, le consentement, les accords entre institutions et l'intégration de l'échange dans le flux de travail.

Un exemple montre pourquoi le niveau syntaxique ne suffit pas. Une étude a suivi 9 analyses de laboratoire transmises entre deux institutions par des messages HL7 v2. Bien que l'échange ait utilisé une messagerie standard, des informations importantes pour la sécurité ont été perdues : la source de l'échantillon manquait systématiquement, les unités et les intervalles de référence étaient incohérents, et les codes LOINC étaient utilisés différemment pour les mêmes tests [299].

8.2.2 HL7 version 2 et CDA

HL7 v2 est le standard de messagerie le plus utilisé dans les hôpitaux. L'organisation HL7 International le présente comme probablement le standard le plus implémenté dans le domaine de la santé, utilisé par 95 % des organisations de santé aux États-Unis et implémenté dans plus de 35 pays [300]. Un message v2 part d'un événement déclencheur (par exemple l'admission d'un patient) et se compose de segments. Chaque segment contient des champs, et les champs contiennent des composants d'un type de données déterminé. Le message identifie l'émetteur, le destinataire, le type de message et le patient [15]. Voici un fragment d'un message d'admission, avec des données fictives :

```
MSH|^~\&|HIS|SPITAL_DEMO|LIS|LAB_DEMO|20260927103000||ADT^A01|MSG00001|P|2.5.1
PID|1||123456^^^SPITAL_DEMO^MR||POPESCU^ION^^^^^L||19800115|M|||STR. EXEMPLU
1^^TIMISOARA^^300000^ROU
```

La barre verticale sépare les champs, et l'accent circonflexe (^) sépare les composants. Le segment MSH (l'en-tête) indique qu'il s'agit d'un message ADT^A01 (admission) en version 2.5.1. Dans le segment PID, le champ 3 contient l'identifiant 123456 attribué par SPITAL_DEMO, de type MR (numéro de dossier), le champ 5 le nom (nom de famille, prénom, type L pour le nom légal), le champ 7 la date de naissance au format AAAAMMJJ, le champ 8 le sexe et le champ 11 l'adresse. Le format est compact, mais de nombreux champs sont facultatifs, et les variantes locales font que chaque interface v2 exige beaucoup de tests.

L'architecture de document clinique (*Clinical Document Architecture*, CDA) Release 2 résout un autre problème : l'échange de documents entiers, comme la lettre de sortie ou le compte rendu d'imagerie. Un document CDA est écrit en XML et contient à la fois une présentation lisible par l'humain et des structures de données exploitables par l'ordinateur. HL7 précise que les développements futurs dans le domaine des documents cliniques structurés devraient se concentrer principalement sur FHIR [301].

8.2.3 openEHR et ISO 13606

Les standards de messagerie décrivent la manière dont circulent les données. Deux autres spécifications décrivent la manière dont elles sont conservées dans le dossier. openEHR est un ensemble ouvert de spécifications, de modèles cliniques et de logiciels qui sépare la partie technique (la manière dont les données sont stockées, décrite par le modèle de référence) de la partie clinique (ce que disent les données), décrite au moyen d'archétypes [15,302]. Les archétypes sont des modèles réutilisables d'un concept clinique, tel que la pression artérielle, construits par des professionnels de la santé et des technologies médicales. Le catalogue international d'archétypes se trouve dans l'outil Clinical Knowledge Manager [302,303]. La spécification couvre l'enregistrement, le stockage et l'interrogation des données [303]. La norme ISO 13606 repose elle aussi sur un modèle de référence et sur des archétypes, mais elle est orientée vers l'échange d'extraits de dossier avec préservation de leur sens complet, et elle est recommandée pour les projets d'interopérabilité régionaux, nationaux ou internationaux [303]. Un groupe d'experts a conclu en 2023 qu'openEHR, ISO 13606 et FHIR doivent être appliqués chacun aux besoins pour lesquels ils ont été créés, et qu'ils peuvent fonctionner ensemble [303]. En pratique, l'adoption reste inégale : dans une revue de 99 études de 2021-2024, 57 % utilisaient le modèle OMOP pour la recherche, 39 % FHIR et seulement 8 % openEHR [304].

8.2.4 HL7 FHIR

FHIR (*Fast Healthcare Interoperability Resources*) a été présenté pour la première fois par HL7 en 2011, afin de dépasser la complexité du développement fondé sur le modèle d'information de référence de HL7 (RIM) ; il a été conçu pour rester compatible avec les systèmes qui utilisent des messages v2 ou v3 et avec les implémentations CDA [15]. L'unité de base est la ressource : tout contenu échangeable est défini comme une ressource, avec des éléments, des types de données et une section lisible par l'humain. Les ressources sont représentées en JSON ou en XML et sont liées entre elles par des références [305]. Cinq ressources couvrent une bonne partie de l'activité clinique quotidienne :

- **Patient**, les données démographiques et les identifiants ;
- **Encounter**, le contact du patient avec les services de santé (consultation, hospitalisation) ;
- **Observation**, les mesures et les résultats (signes vitaux, analyses, scores) ;
- **Condition**, les problèmes et les diagnostics ;
- **MedicationRequest**, la prescription d'un médicament.

L'accès se fait par une interface REST, c'est-à-dire par les requêtes habituelles du web. `GET [base]/Patient/123` lit le patient dont l'identifiant est 123, `GET [base]/Observation?patient=123` recherche toutes ses observations et renvoie un paquet (Bundle), et `POST [base]/Observation` crée une nouvelle observation (figure 8.3). Voici une ressource Patient minimale, fictive :

```
{
  "resourceType": "Patient",
  "id": "123",
  "identifier": [{ "system": "urn:oid:1.2.3.4.5", "value": "1800115123456" }],
  "name": [{ "family": "Popescu", "given": ["Ion"] }],
  "gender": "male",
  "birthDate": "1980-01-15"
}
```

Les mêmes données que dans le segment PID apparaissent ici avec des noms de champs explicites. L'identifiant et son système (`system`) sont fictifs.

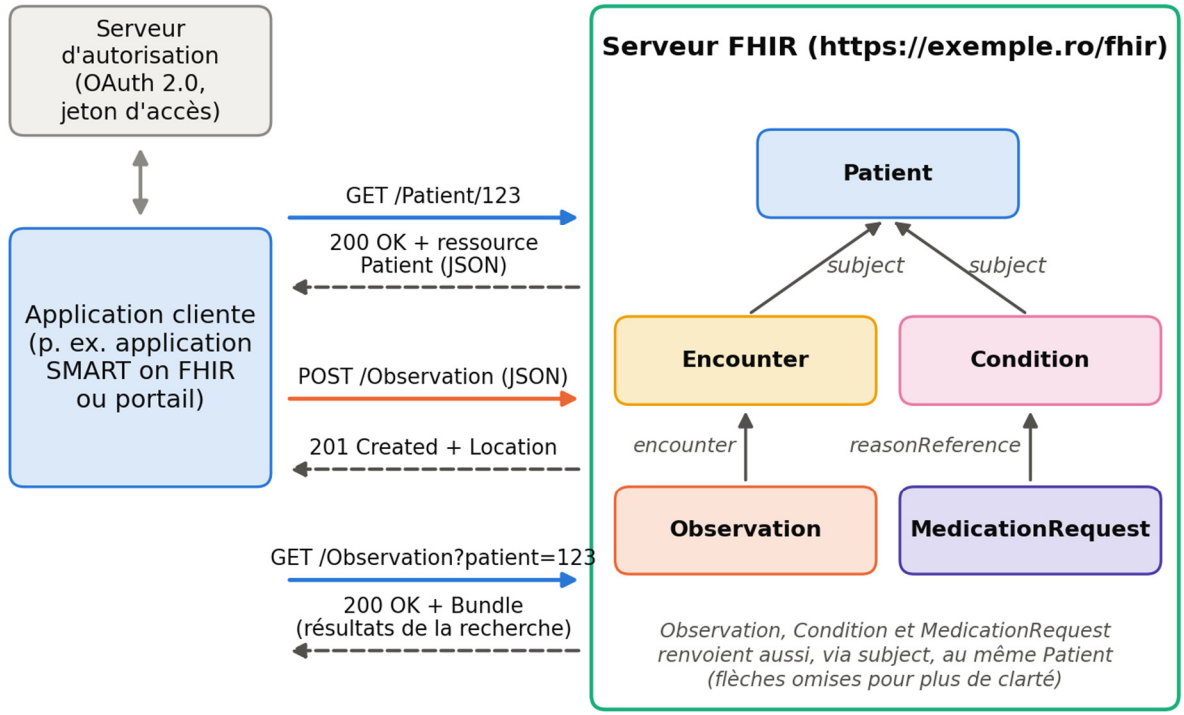


Figure 8.3. Interaction REST entre une application cliente et un serveur FHIR (lecture, création, recherche) et graphe des références entre les ressources Patient, Encounter, Observation, Condition et MedicationRequest. Les noms des éléments de liaison sont ceux de FHIR R4. Adapté de [305,306].

Les ressources de base sont générales. Un profil restreint une ressource pour un contexte donné, par exemple en fixant les terminologies obligatoires ou les champs à renseigner, et un guide d'implémentation (*implementation guide*, IG) rassemble les profils, les règles et les exemples relatifs à un cas d'utilisation, comme le résumé du patient de la section 8.3. Le mécanisme d'extensions permet d'ajouter des éléments sans modifier les ressources de base [305].

FHIR est publié en versions successives. La Release 4 (4.0.1, 30 octobre 2019) a apporté le premier contenu normatif, stable, et c'est sur elle que s'appuient la plupart des implémentations et des guides actuels. R4B (4.3.0, 28 mai 2022) a été une version intermédiaire, et R5 (5.0.0, 26 mars 2023) est la dernière version publiée. R6 n'était pas publiée au 27 septembre 2026 : le deuxième vote complet (6.0.0-ballot5) et un instantané étiqueté "stable" (6.0.0-snapshot1) sont

parus le 17 juillet 2026, tous deux avec le statut de projet [307]. Pour un étudiant, la règle pratique est simple : les guides d'implémentation cités dans ce chapitre (IPS, SMART App Launch, CDS Hooks) et le serveur public de test du TP 8 utilisent R4.

La revue des 99 études citée plus haut a observé que les implémentations FHIR nécessitent souvent des adaptations locales, ce qui complique précisément l'interopérabilité recherchée [304].

8.2.5 SMART on FHIR et CDS Hooks

SMART App Launch est le guide HL7 qui décrit comment une application externe obtient une autorisation, s'authentifie et s'intègre à un serveur FHIR. Il repose sur OAuth 2.0 et définit deux scénarios : le lancement depuis le DSE, lorsque l'application reçoit le contexte (par exemple le patient ouvert à ce moment-là), et le lancement autonome, lorsque l'utilisateur choisit le patient après le démarrage. Les autorisations s'expriment par des domaines d'accès (*scopes*), par exemple le droit de lire et de rechercher les ressources Encounter. La version actuelle est la STU 2.2, publiée le 30 avril 2024 [306]. Sur cette base, un groupe de chercheurs a intégré deux applications de résultats rapportés par les patients dans 18 cabinets qui utilisaient 3 plateformes de DSE différentes [308].

CDS Hooks est complémentaire. La spécification décrit comment un DSE appelle, par une API REST et des messages JSON, un service externe d'aide à la décision à certains moments du flux, appelés *hooks*, par exemple l'ouverture du dossier (`patient-view`) ou la signature d'une prescription (`order-sign`). Le service répond par des "cartes" affichées au clinicien [309]. La version actuelle est la STU 2 (2.0.1, mars 2025), et une version normative (R3) a été soumise au vote en décembre 2025 [309]. Dans un essai randomisé mené auprès de 70 cliniciens d'un service des urgences, les cartes CDS Hooks qui suggéraient un calculateur clinique ont fait passer l'utilisation de l'application SMART correspondante de 2,6 % à 6,0 % des interactions (OR 2,45) [310].

Tableau 8.2. Les principaux standards et normes du dossier électronique et leur rôle. Source : [300-303,305,306,309,311,312].

Standard / norme	Organisation	Niveau	Rôle
HL7 v2	HL7 International	syntaxique	messages entre les systèmes d'un hôpital (admissions, laboratoire)
CDA R2	HL7 International	syntaxique	documents cliniques en XML (lettre de sortie, compte rendu)
FHIR R4/R5	HL7 International	syntaxique, avec des liens sémantiques	ressources et API REST pour les applications et l'échange
openEHR	openEHR International	sémantique (modélisation)	stockage du dossier fondé sur les archétypes
ISO 13606	ISO/CEN	sémantique (modélisation)	communication d'extraits de dossier
ISO 27269	ISO/CEN	contenu	jeu de données du résumé international du patient
ISO 17523	ISO/CEN	contenu	exigences relatives aux données d'une prescription électronique
SMART App Launch	HL7 International	technique, sécurité	autorisation des applications sur des données FHIR
CDS Hooks	HL7 International	intégration dans le flux	appel de services d'aide à la décision

8.3 Le résumé international du patient et les services transfrontaliers

8.3.1 IPS

Un médecin de garde qui reçoit un touriste inconscient a besoin de quelques informations, et non du dossier entier : problèmes actifs, allergies, médicaments. Le résumé international du patient (*International Patient Summary*, IPS) répond à ce besoin. La norme ISO 27269 définit le jeu de données de l'IPS comme minimal et non exhaustif, destiné aux soins programmés et non programmés au-delà de toute frontière organisationnelle, y compris entre pays [311]. Le guide HL7 qui la met en œuvre reprend de la norme la précision selon laquelle ce jeu de données est indépendant de la spécialité et de la maladie [313]. Ce jeu de données a d'abord fait l'objet d'une norme européenne en 2019, promue ensuite au rang de norme internationale en 2021. La deuxième édition, ISO 27269:2025, reste une définition abstraite, dont on dérive des modèles implémentables. Elle ajoute les sections "Alertes" et "Récit du patient", tandis que les anciennes sections "Cross-Border" et "Provenance" sont fusionnées en un seul bloc, "Cross Organization" [311].

La principale implémentation technique est le guide HL7 FHIR IPS. La version actuelle, 2.0.1 (STU 2), a été publiée le 19 juin 2026 et repose sur FHIR R4 [313]. Le guide divise le résumé en sections. Trois sont obligatoires : la liste des problèmes, les allergies et intolérances, et le résumé des médicaments. Les vaccinations, les résultats d'examens, les procédures et les dispositifs médicaux sont recommandés, et le reste (signes vitaux, histoire sociale, grossesse, directives anticipées) est facultatif [313]. Le résumé peut aussi circuler par l'intermédiaire du patient : quatre implémentations aux États-Unis, au Canada et dans l'organisation du pèlerinage du Hajj ont utilisé l'IPS avec des codes QR et des SMART Health Links, afin que le patient présente lui-même ses antécédents [314].

La limite de l'IPS tient justement à sa concision. Une analyse réalisée pour deux projets européens consacrés à l'insuffisance cardiaque a montré que le guide IPS peut aussi servir à la recherche, mais seulement après des extensions (mineures, selon les auteurs) comportant des éléments spécifiques [315].

8.3.2 La prescription électronique

La norme ISO 17523, dans sa deuxième édition de 2025, fixe les exigences relatives au contenu d'une prescription électronique, c'est-à-dire du document numérique émis par le médecin prescripteur et reçu par le pharmacien qui délivre le médicament. Elle ne réglemente ni le mode de transmission, ni la validation, ni le remboursement, laissés à l'appréciation de chaque pays, et s'appuie sur les normes ISO d'identification des médicaments (IDMP) [312]. Deux systèmes nationaux conformes à l'ISO 17523 peuvent donc rester techniquement incompatibles en l'absence d'une infrastructure d'échange commune.

8.3.3 MyHealth@EU

L'infrastructure européenne commune s'appelle MyHealth@EU. Elle offre deux services : la prescription et la dispensation électroniques, grâce auxquelles un citoyen peut retirer ses médicaments dans une pharmacie d'un autre État membre, et le résumé du patient, qui met à la disposition du médecin du pays de traitement les informations de base sur le patient, dans sa propre langue [316]. Chaque État se connecte par l'intermédiaire d'un point de contact national, et les données circulent entre ces points (figure 8.4) [23].

L'utilisation réelle a progressé lentement. Une analyse des indicateurs officiels de 2019 au deuxième trimestre de 2022 a dénombré onze pays disposant d'au moins un service opérationnel, dont neuf avaient déclaré des transactions. Le programme servait encore surtout à des tests, et la dispensation transfrontalière de médicaments concernait principalement des patients estoniens et finlandais [317]. Dans une enquête menée auprès des pharmaciens de ces deux pays, 88 % des Estoniens et 81,8 % des Finlandais ont estimé que la prescription électronique transfrontalière avait amélioré l'accès aux médicaments. Par ailleurs, 61 % et 42,8 % d'entre eux, respectivement, avaient rencontré des ambiguïtés ou des erreurs dans les prescriptions [318].

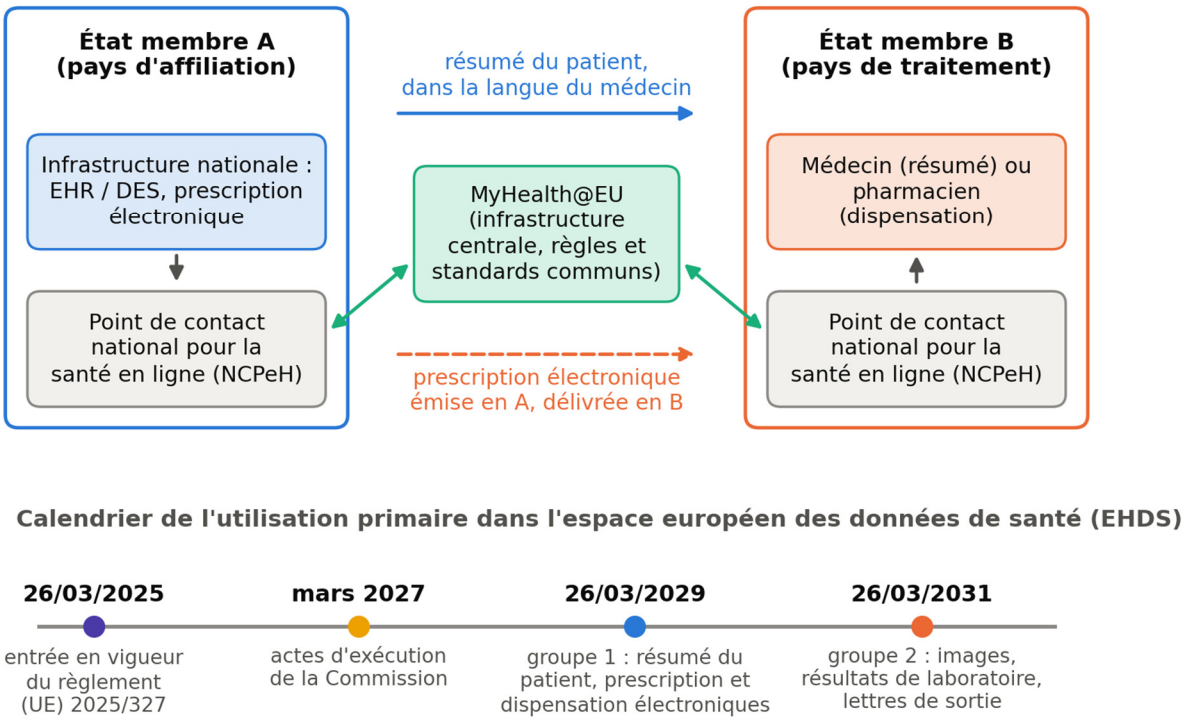


Figure 8.4. L'échange transfrontalier par MyHealth@EU : le résumé du patient et la prescription électronique circulent entre les points de contact nationaux de l'État d'affiliation (A) et de l'État de traitement (B). En bas, le calendrier de l'utilisation primaire prévu par le règlement (UE) 2025/327. Adapté de [316,319].

8.4 L'espace européen des données de santé

Le règlement (UE) 2025/327 relatif à l'espace européen des données de santé (*European Health Data Space*, EHDS) a été adopté le 11 février 2025 et publié au Journal officiel le 5 mars 2025 ; il est entré en vigueur le 26 mars 2025. Il s'applique, en règle générale, à partir du 26 mars 2027, mais les obligations concrètes entrent en vigueur de manière échelonnée [23,319]. Le règlement distingue deux utilisations des données. L'utilisation primaire désigne l'emploi des données pour les soins de la personne concernée, et l'utilisation secondaire leur réutilisation à des fins de recherche, d'innovation, de politiques publiques ou d'activités réglementaires [319,320].

8.4.1 L'utilisation primaire

Pour l'utilisation primaire, le règlement définit six catégories prioritaires de données : les résumés des patients, les prescriptions électroniques, les dispensations électroniques, les études d'imagerie médicale et les comptes rendus associés, les résultats des examens médicaux (y compris de laboratoire) et les comptes rendus associés, ainsi que les rapports de sortie [23]. Le calendrier comporte deux étapes (figure 8.4). À partir du 26 mars 2029, les droits des patients et

les obligations liées aux trois premières catégories, ainsi qu'aux systèmes de DSE qui les traitent, s'appliquent dans tous les États membres. À partir du 26 mars 2031 s'y ajoutent les images, les résultats de laboratoire et les rapports de sortie [23,319].

Le patient reçoit des droits explicites : l'accès à ses propres données relevant des catégories prioritaires et le droit de restreindre l'accès des professionnels à certaines données. Les États membres peuvent en outre prévoir un droit d'opposition (*opt-out*) à l'utilisation primaire [23]. Chaque État désigne un point de contact national pour la santé numérique, connecté à MyHealth@EU, de sorte que l'échange transfrontalier décrit à la section 8.3 devient une obligation, et non une option [23]. Le format technique commun, appelé format européen d'échange des dossiers de santé électroniques (*European Electronic Health Record Exchange Format*, EEHRxF), s'appuie sur des spécifications telles que le résumé du patient décrit plus haut [315].

8.4.2 Exigences applicables aux systèmes de DSE

L'EHDS réglemente aussi les logiciels. Un système de DSE qui traite des catégories prioritaires doit comporter deux composants logiciels harmonisés : le composant européen d'interopérabilité et le composant européen de journalisation, qui enregistre qui a accédé aux données. Le système ne peut être mis sur le marché que s'il respecte les exigences du chapitre consacré aux systèmes de DSE [23]. Un hôpital qui achète aujourd'hui un logiciel doit donc vérifier ces exigences.

8.4.3 L'utilisation secondaire, en bref

Le chapitre relatif à l'utilisation secondaire s'applique lui aussi à partir du 26 mars 2029, certaines catégories (par exemple les données génétiques et génomiques) n'étant concernées qu'à partir de 2031. L'accès des chercheurs se fait sur la base d'une autorisation délivrée par un organisme responsable de l'accès aux données de santé, et l'échange transfrontalier passe par l'infrastructure HealthData@EU [23,319]. Une revue rapide (*rapid review*) de la littérature consacrée aux implications éthiques et sociales de l'EHDS a constaté que la plupart des problèmes discutés concernent l'utilisation secondaire et que les États qui rencontrent déjà des difficultés de numérisation risquent d'être touchés de manière disproportionnée [320].

8.5 La Roumanie en 2026

La description ci-dessous reflète les sources officielles disponibles au 27 septembre 2026. L'infrastructure est en pleine mutation.

8.5.1 Le dossier électronique de santé roumain (DES)

Le fondement juridique du dossier électronique de santé roumain (Dosarul electronic de sănătate, DES) est le titre IX^1 de la loi n° 95/2006, introduit par la loi n° 45/2019 (articles 346^1-346^12). La loi décrit le DES comme un ensemble d'enregistrements électroniques consolidés au niveau national, comportant des données cliniques, biologiques, diagnostiques et thérapeutiques. Le médecin accède au dossier avec le consentement du patient. L'exception est le module "Sumar de urgență" (résumé d'urgence), accessible sans consentement aux médecins des structures d'urgence (salle de garde, services d'urgence UPU et CPU, ambulance) et des soins primaires. Le patient peut refuser expressément l'utilisation du DES ; ses données sont alors anonymisées, sans qu'il perde l'accès aux services médicaux [321]. Dans les termes de ce chapitre, le DES est un DSE national.

8.5.2 La plateforme informatique de l'assurance maladie

Le DES fait partie de la Platforma informatică din asigurările de sănătate (PIAS, plateforme informatique de l'assurance maladie), aux côtés du Sistemul informatic unic integrat (SIUI, système informatique unique intégré), du Sistemul informatic de prescripție electronică (SIPE, système informatique de prescription électronique) et du cardul electronic de asigurări de sănătate (CEAS, carte électronique d'assurance maladie) [322]. Le projet PNRR (plan national de relance et de résilience) "Redimensionarea, standardizarea și optimizarea Platformei Informatice din Asigurările de Sănătate - PIAS" (redimensionnement, standardisation et optimisation de la plateforme informatique de l'assurance maladie) a été financé au titre de la composante 7 (Transformation numérique), investissement I3.1, et a été déclaré achevé le 31 août 2026. La valeur approuvée était de 493 900 milliers de lei hors TVA, et le montant payé, de 434 679 milliers de lei hors TVA. Le communiqué d'achèvement énumère la nouvelle infrastructure, la migration de la plateforme existante, le portail e-Sănătatea Mea, les formulaires médicaux électroniques et le registre des assurés [323]. La formulation officielle est "PIAS modernizată" (PIAS modernisée), c'est-à-dire la même plateforme, redimensionnée et optimisée, et non un système portant un nouveau nom [323].

Pour le patient, le résultat visible est le portail eSănătateaMea, lancé le 1er septembre 2026. L'accès se fait par ROeID, le système national d'identité numérique, et le portail affiche la qualité d'assuré, les ordonnances émises au cours des 5 dernières années, l'historique de santé des 10 dernières années, l'historique des remboursements, la situation de la carte de santé et les rendez-vous en ligne chez les prestataires [324]. Dans les termes de la section 8.1, le portail fonctionne comme un PHR à l'échelle nationale. Le point de départ est modeste : en 2024, seulement 9,9 % des personnes en Roumanie avaient consulté en ligne leur dossier de santé (figure 8.1) [297].

8.5.3 La carte de santé et la prescription électronique

La carte nationale d'assurance maladie reste en usage. La CNAS (Caisse nationale d'assurance maladie) a précisé qu'à partir du 1er septembre 2026 l'utilisation de la carte d'identité électronique ne devenait pas obligatoire pour les prestataires et que l'accès des assurés aux services n'en dépendait pas [325]. La prescription électronique fonctionne au moyen du SIPE, composant de la PIAS, et les ordonnances émises parviennent désormais aussi au portail du patient [322,324]. La différence par rapport au modèle européen décrit à la section 8.3 relève des niveaux organisationnel et sémantique : pour qu'une ordonnance roumaine puisse être délivrée dans une pharmacie d'un autre État, le SIPE doit être connecté à MyHealth@EU par l'intermédiaire du point de contact national, obligation que l'EHDS fixe pour mars 2029 [23]. Un obstacle sémantique concret est la terminologie : au 27 septembre 2026, la Roumanie ne figurait pas parmi les membres de SNOMED International [151].

8.5.4 La numérisation des hôpitaux dans le cadre du PNRR

Toujours dans la composante 7, l'investissement I3.3 a financé des systèmes informatiques et une infrastructure numérique pour environ 200 établissements de santé publics, avec un budget de 493 900 000 lei hors TVA [326]. À la date de rédaction, les sources officielles consultées ne présentaient pas les résultats physiques agrégés de ces projets.

Étude de cas

- L'hôpital clinique départemental d'urgence "Pius Brînzeu" de Timişoara (Spitalul Clinic Judeţean de Urgenţă "Pius Brînzeu" Timişoara) a annoncé, le 17 août 2024, le projet "Sănătatea, siguranța și confortul pacientului prin servicii integrate digitale la SCJU Timișoara" (santé, sécurité et confort du patient grâce à des services numériques intégrés au SCJU Timișoara), code PNRR/C7/MS-733/I3.3, d'une valeur totale de 5 877 410,00 lei TVA incluse [327].
- Activités déclarées : équipements de stockage pour la migration des données, terminaux (ordinateurs de bureau, tablettes médicales, lecteurs de codes-barres et de codes QR), applications mobiles pour les médecins et les infirmières destinées à faciliter l'intégration des données du dossier électronique du patient, un logiciel d'intelligence artificielle pour l'analyse des TDM pulmonaires, applications pour le circuit des consommables médicaux et des médicaments, formation du personnel [327].
- L'objectif déclaré comprend l'amélioration de l'interopérabilité et des échanges de données au sein de l'hôpital, avec les fournisseurs, avec les patients et avec d'autres entités [327]. La page du projet, consultée le 27 septembre 2026, ne mentionne ni la date d'achèvement ni les résultats obtenus.
- Questions pour la discussion : quelles activités relèvent du niveau technique et lesquelles du niveau sémantique ? Que faut-il vérifier d'ici 2029, lorsque les résumés des patients deviendront obligatoires en vertu de l'EHDS ?

8.6 Utilisabilité, sécurité et facteurs humains

Un DSE peut produire des erreurs du fait de la manière dont il est construit et utilisé. Une analyse de 2 700 rapports d'incidents provenant de plus de 595 établissements de santé a relevé 1 508 erreurs médicamenteuses associées aux systèmes informatiques, dont la moitié avaient atteint le patient. Presque toutes (97,3 %) impliquaient un problème d'utilisabilité, en particulier la saisie des données, le soutien au flux de travail et les alertes, et 80,5 % étaient des erreurs de dose [328]. Une enquête nationale menée auprès de 1 933 médecins a attribué 38 % de la variation des scores d'utilisabilité aux différences entre produits et 51 % aux différences entre hôpitaux utilisant le même produit [329]. Le mode d'implémentation semble compter autant que le logiciel.

8.6.1 La fatigue face aux alertes

Les alertes d'aide à la décision ne sont utiles que si elles sont rares et pertinentes. La fatigue face aux alertes (*alert fatigue*) désigne le fait d'ignorer les alertes en raison du grand nombre d'avertissements non pertinents. Dans un centre médical universitaire, 95,7 % des 16 011 alertes d'interactions médicamenteuses de haute priorité ont été ignorées, y compris 87,3 % de celles de gravité maximale. Seules 45,4 % des alertes ignorées l'ont été à juste titre, et les événements indésirables ont été plus fréquents lorsque l'alerte avait été ignorée à tort (9,4 % contre 4,3 %) [330]. La mesure du phénomène n'est pas encore standardisée : sur 22 revues systématiques analysées en 2026, une seule définissait de manière opérationnelle la fatigue face aux alertes [331]. Dans l'enquête auprès des médecins de famille citée plus haut, l'utilité des alertes a obtenu la note la plus faible [295].

8.6.2 Le copier-coller et les notes gonflées

La facilité avec laquelle on copie du texte d'une note à l'autre fait gagner du temps, mais produit des "notes gonflées" (*note bloat*) et peut propager des informations qui ne sont plus vraies, par exemple un "hier" copié dans une note écrite un autre jour [285]. L'analyse de 104 millions de notes d'un système universitaire américain a montré que 50,1 % du texte total était dupliqué à

partir de notes antérieures concernant le même patient, et que cette proportion était passée de 33,0 % en 2015 à 54,2 % en 2020. Près de la moitié du texte dupliqué provenait d'un autre auteur [332].

8.6.3 Mesures de sécurité

Les guides SAFER, utilisés par les hôpitaux américains pour l'autoévaluation de la sécurité du DSE, ont été révisés en 2025 : le nombre de pratiques recommandées a diminué de 40 %, et chaque recommandation a reçu un niveau de preuve [333]. Indépendamment des guides, tout DSE peut connaître des interruptions, planifiées ou non ; c'est pourquoi chaque institution a besoin de procédures de travail pour les périodes pendant lesquelles le système n'est pas disponible [285]. Pour le clinicien, quelques règles simples réduisent le risque : vérifier l'identité du patient avant chaque prescription, relire le texte copié, signaler les alertes inutiles à l'équipe qui paramètre le système.

Pièges fréquents

- Croire qu'un système "compatible HL7" communique automatiquement avec n'importe quel autre système. Sans profils ni terminologies communs, les messages peuvent être formellement corrects et néanmoins mal compris.
- Confondre le DSE et le PHR. Le DSE est administré par les prestataires de soins, le PHR par le patient.
- Considérer FHIR comme une base de données ou un produit. FHIR est un standard (ressources et API), que mettent en œuvre différents serveurs et applications.
- Traiter le résumé du patient comme une copie du dossier entier. L'IPS est volontairement minimal.
- Ignorer toutes les alertes par réflexe. Parmi les alertes non pertinentes se trouvent aussi des alertes qui préviennent des événements indésirables réels.

À retenir

- Le DSE est un système actif, et non une archive : il vérifie, affiche des tendances, envoie des alertes.
- Les données structurées sont réutilisables, le texte libre conserve le raisonnement. La plupart des systèmes combinent les deux.
- Les médecins passent une grande partie de leur journée dans le DSE, et ce temps augmente. L'épuisement professionnel est associé à cette charge.
- L'interopérabilité comporte quatre niveaux. Le niveau sémantique exige des terminologies communes.
- HL7 v2 domine dans les hôpitaux, FHIR domine dans les nouvelles applications et dans l'échange par API.
- L'IPS comporte trois sections obligatoires : problèmes, allergies, médicaments.
- L'EHDS oblige les États membres à échanger le résumé du patient et les prescriptions électroniques à partir de mars 2029.

Questions d'autoévaluation

1. (QCS) Quel niveau de l'interopérabilité est en cause lorsque deux laboratoires envoient le même test correctement formaté, mais avec des codes LOINC différents ?

a) technique
b) syntaxique
c) sémantique
d) organisationnel
e) physique

2. (QCS) Dans le segment HL7 v2 `PID|1||123456^^^SPITAL_DEMO^MR||POPESCU^ION^^^^^L||19800115|M`, la valeur 19800115 représente :

a) l'identifiant du patient
b) la date d'admission
c) la date de naissance
d) le code du service
e) le numéro du message

3. (QCS) Laquelle des sections suivantes est obligatoire dans le guide HL7 FHIR IPS ?

a) les signes vitaux
b) l'histoire sociale
c) les vaccinations
d) les allergies et intolérances
e) les directives anticipées

4. (QCM) À propos de HL7 FHIR, les affirmations vraies sont les suivantes :

a) l'unité de base de l'échange est la ressource
b) les ressources peuvent être représentées en JSON ou en XML
c) la recherche des observations d'un patient se fait par une requête GET
d) FHIR impose une base de données centrale unique pour tous les hôpitaux
e) les profils restreignent les ressources pour un cas d'utilisation donné

5. (QCM) Concernant la charge de documentation dans le DSE, les études citées dans le chapitre montrent que :

a) les médecins passent dans le DSE environ un tiers de leur journée de travail
b) les cliniciens des États-Unis travaillent activement plus longtemps dans le DSE que ceux qui exercent hors des États-Unis
c) les scribes fondés sur l'intelligence artificielle ont complètement supprimé le travail après les heures de service
d) le temps de travail dans le DSE après les heures de service a été associé à une probabilité plus élevée d'épuisement
e) le temps passé dans le DSE a diminué de façon constante après 2019

CHAPITRE 9

Chapitre 9. La télémédecine : formes, preuves, limites et cadre juridique

Objectifs

À la fin de ce chapitre, l'étudiant sera capable :

1. de définir la télémédecine, la télésanté et la e-santé et de rattacher un service donné à l'une des formes de la télémédecine (téléconsultation, téléexpertise, télésurveillance, téléassistance), en précisant les acteurs et la modalité synchrone ou asynchrone.
2. de décrire les composants techniques d'un système de téléconsultation ou de télésurveillance, y compris les couches de sécurité et le lien avec le dossier de santé électronique.
3. d'interpréter les résultats des principales revues systématiques publiées en 2020-2026 pour l'insuffisance cardiaque, l'hypertension, le diabète, la psychiatrie, la dermatologie, l'accident vasculaire cérébral et la radiologie.
4. d'énumérer les limites et les risques de la téléconsultation et de reconnaître les situations qui imposent un examen en présentiel ou un appel d'urgence.
5. d'expliquer le cadre juridique et le régime de remboursement de la télémédecine en Roumanie au 27 septembre 2026, ainsi que les repères européens applicables.
6. de rédiger une note de téléconsultation complète à partir d'un modèle.

9.1 Définitions et taxonomie

9.1.1 Télémédecine, télésanté, e-santé

L'Organisation mondiale de la santé (OMS) définit la télémédecine comme la prestation de services de santé, là où la distance est un facteur critique, par tout professionnel de santé utilisant les technologies de l'information et de la communication (TIC) pour échanger des informations valides à des fins de diagnostic, de traitement et de prévention [334]. La définition comporte trois éléments : la distance, le professionnel et l'information valide. Un message envoyé à un ami médecin sur une application de messagerie ne remplit pas automatiquement ces conditions.

La télésanté (*telehealth*) est une notion plus large. Elle inclut la télémédecine, mais aussi la formation à distance des professionnels et d'autres applications dans lesquelles les communications électroniques soutiennent les services de santé [334]. La e-santé (*eHealth*) recouvre tout usage des TIC en santé : le dossier électronique, l'ordonnance électronique, les registres, les applications mobiles. Les trois termes sont liés par une relation d'inclusion (figure 9.1) : tout acte de télémédecine est un acte de e-santé, mais la réciproque n'est pas vraie.

9.1.2 Modalités : synchrone, asynchrone, surveillance continue

Le principal critère technique est le moment où l'autre participant répond. Le guide de l'OMS de 2022 décrit trois modes [334] :

- **synchrone** (interactif, en temps réel) : le patient et le médecin, ou deux médecins, communiquent simultanément, par vidéo ou par téléphone.
- **asynchrone**, appelé aussi stockage et transmission (*store-and-forward*) : les données et les images sont envoyées puis analysées plus tard par un spécialiste, en général dans des situations non urgentes (dermatologie, radiologie, ophtalmologie).
- **surveillance à distance** (télésurveillance) : des dispositifs médicaux connectés transmettent des paramètres, et l'équipe médicale reçoit des alertes lorsque les valeurs sortent des intervalles fixés.

Le second critère tient aux acteurs. La télémédecine peut relier le patient au professionnel (*client-to-provider*) ou deux professionnels entre eux (*provider-to-provider*). Le guide de l'OMS cite parmi les usages fréquents la téléconsultation (services cliniques à distance pour le patient), la téléexpertise (un médecin demande un second avis à un spécialiste), le télétriage (la hiérarchisation des décisions d'urgence à partir des paramètres transmis) et l'auto-assistance guidée, utilisée en santé mentale [334]. La téléassistance désigne le soutien à distance des personnes vivant à domicile (bouton d'alarme, appel programmé, accompagnement de l'aidant). Sa définition juridique en Roumanie est examinée à la section 9.5.

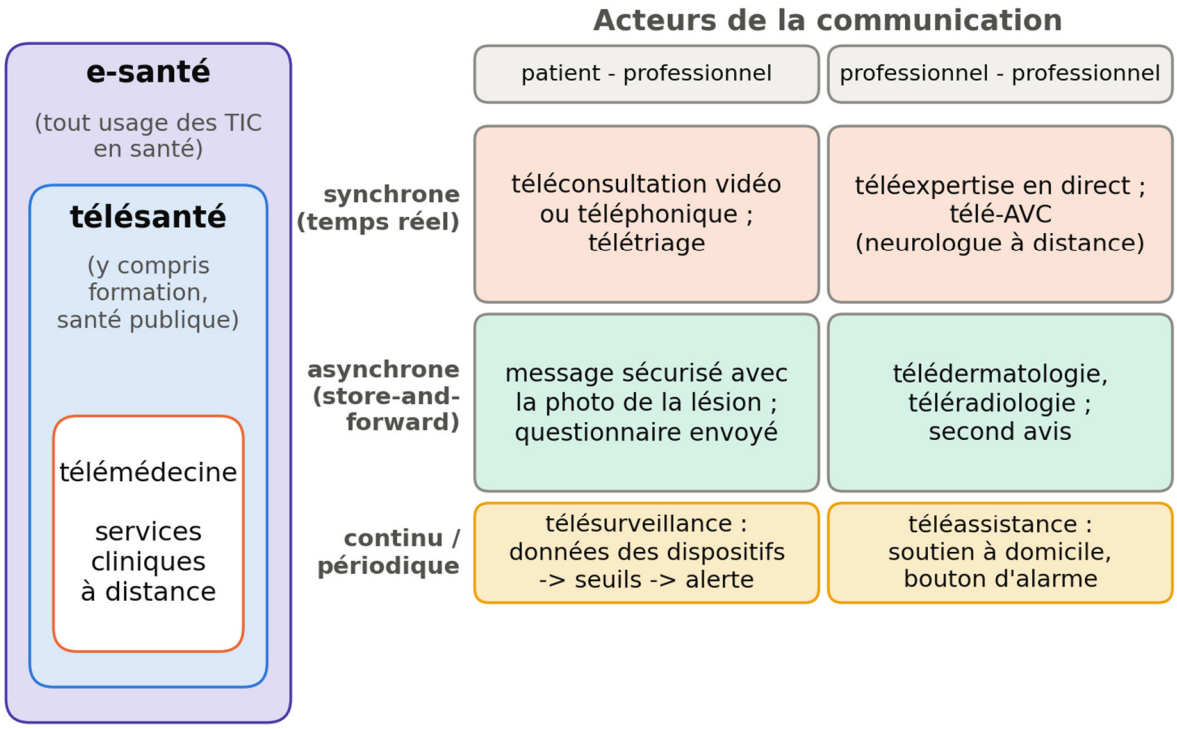


Figure 9.1. Taxonomie de la télémédecine : relation d'inclusion entre e-santé, télésanté et télémédecine (à gauche) et formes de la télémédecine selon les acteurs et la modalité (à droite). Élaboré d'après [334,335].

Tableau 9.1. Formes de la télémédecine, avec exemples. Source : [334,335].

Forme	Acteurs	Modalité	Exemple	Code OMS 2023
Téléconsultation	patient - médecin	synchrone (vidéo, téléphone)	contrôle à 3 mois pour une hypertension	2.4.1
Téléexpertise	médecin - médecin	synchrone ou asynchrone	le médecin de famille envoie au dermatologue une image dermatoscopique	2.4.4
Télésurveillance	dispositif - équipe	continue ou périodique	poids et pression artérielle transmis chaque jour dans l'insuffisance cardiaque	2.4.2
Transmission de données médicales	patient ou médecin - spécialiste	asynchrone	radiographie envoyée pour interprétation (téléradiologie)	2.4.3
Téléassistance	patient, aidant - service	synchrone, à la demande	bouton d'alarme pour la personne âgée vivant seule	pas de code distinct

9.1.3 La classification OMS des interventions numériques (2e édition, 2023)

Pour décrire un projet sans ambiguïté, l'OMS a publié en 2023 la deuxième édition de la classification des interventions, services et applications de santé numérique [335]. La classification comporte trois axes : les problèmes du système de santé (par exemple, "prestation tardive des soins"), les interventions numériques et les types de services et d'applications. Les interventions sont regroupées selon leur utilisateur principal : les personnes (groupe 1), les prestataires de soins de santé (groupe 2), le personnel de gestion et d'appui (groupe 3), les services de données (groupe 4).

La télémédecine possède un sous-groupe propre, 2.4, qui compte quatre interventions : 2.4.1 consultations entre la personne à distance et le prestataire, 2.4.2 surveillance à distance, par le prestataire, des données de santé ou de diagnostic, 2.4.3 transmission de données médicales (images, notes, vidéos) au prestataire, 2.4.4 consultations entre prestataires pour la prise en charge du cas [335]. Pour 2.4.1, la classification donne comme synonymes la télémédecine interactive et synchrone, et pour 2.4.3 le stockage et la transmission, c'est-à-dire la télémédecine asynchrone. Une distinction utile pour l'étudiant : l'autosurveillance réalisée par le patient, sans que les données parviennent au médecin, est codée à part (1.4.2), dans le groupe des personnes. Un glucomètre qui se contente d'afficher des valeurs ne relève pas de la télésurveillance, bien qu'il s'agisse d'un dispositif numérique.

9.2 L'infrastructure technique

9.2.1 La visioconférence et la bande passante

La téléconsultation vidéo a besoin d'une connexion stable dans les deux sens. Le guide de l'OMS souligne que la télémédecine exige en règle générale une bande passante plus élevée que les autres interventions numériques, si bien que la fracture numérique s'y fait sentir plus fortement [334]. Une étude portant sur 3 659 699 anciens combattants américains a classé la connexion selon les débits déclarés à l'autorité de régulation : inadéquate (débit descendant d'au plus 25 Mbit/s, débit montant d'au plus 3 Mbit/s), adéquate et optimale (au moins 100/100 Mbit/s) [336]. Après le début de la pandémie, les patients disposant d'une connexion optimale ont davantage utilisé les consultations vidéo de santé mentale que ceux dont la connexion était inadéquate (rapport des taux d'incidence 1,52, IC à 95 % 1,45-1,59) [336]. Par conséquent, tout service vidéo doit prévoir une solution de repli (appel téléphonique) et un plan en cas de coupure de la connexion.

9.2.2 Les dispositifs périphériques

L'examen physique à distance s'appuie sur des dispositifs périphériques : stéthoscope numérique, dermatoscope fixé au téléphone, otoscope vidéo, tensiomètre et oxymètre de pouls connectés. Dans une évaluation portant sur 57 patients présentant des souffles cardiaques, la concordance entre l'auscultation à distance au moyen d'un stéthoscope électronique et l'auscultation classique, pour la détection des souffles, a été de 90 % [337]. En télédermatologie, une méta-analyse de 155 études a montré que l'utilisation du dermatoscope a fait passer la concordance diagnostique pour les cancers cutanés de 67 % à 80 % [338].

Tous les signes cliniques ne peuvent pas être transmis. Une revue de portée (*scoping review*) de 25 études de la période pandémique a identifié trois types principaux d'examen physique virtuel (musculo-squelettique, tête et cou, thorax), mais aussi des obstacles : des problèmes techniques, des doutes sur l'efficacité, la confiance de l'infirmier qui manipule le dispositif, l'état du patient et son niveau de littératie en santé [339].

9.2.3 Plateformes et sécurité

Une plateforme de télémédecine réunit la visioconférence, la messagerie sécurisée, le stockage des documents, la prise de rendez-vous et, pour la télésurveillance, les règles d'alerte. Parmi les exigences techniques, le guide de l'OMS mentionne l'authentification des utilisateurs et indique que la surveillance à distance peut nécessiter des solutions dans le cloud avec des données chiffrées [334]. Le même document précise que l'enregistrement audio ou vidéo d'une consultation ne se fait qu'avec le consentement éclairé du patient et que l'archive doit être protégée [334].

La distinction entre le chiffrement "par tronçons" et le chiffrement de bout en bout (*end-to-end encryption*, E2EE) est bien illustrée par la documentation de Jitsi Meet, une plateforme vidéo à code source ouvert. Tout le trafic audio-vidéo est chiffré sur le réseau (DTLS-SRTP), mais ce chiffrement est retiré sur le serveur qui redistribue les flux, où les paquets transitent par la mémoire sans être stockés [340]. Le chiffrement de bout en bout n'est disponible que dans certains navigateurs et couvre le son, l'image et le partage d'écran, mais pas les messages du chat [340]. Conséquence pratique : le médecin doit savoir ce que la plateforme protège et ce qu'elle ne protège pas, et ne pas envoyer de données sensibles par le canal non protégé.

Les données concernant la santé constituent une catégorie particulière de données à caractère personnel (art. 9 du règlement général sur la protection des données, RGPD), et l'art. 32 oblige le responsable du traitement à mettre en œuvre des mesures techniques et organisationnelles adaptées au risque, parmi lesquelles le chiffrement [118]. Une revue des interventions de télésanté de la période pandémique a regroupé 28 difficultés signalées en difficultés techniques (14), non techniques (10) et relatives à la confidentialité et aux politiques (4) ; les problèmes techniques les plus critiques concernaient le réseau, la fiabilité et la compatibilité des systèmes [341].

9.2.4 L'interopérabilité avec le dossier électronique

Un service de télémédecine isolé produit des informations qui n'arrivent pas dans le dossier de santé électronique (DSE, *electronic health record*, EHR) du patient. Le guide de l'OMS recommande d'éviter les déploiements "en silo" en réutilisant des composants communs : identifiants uniques des patients, des professionnels et des établissements, services de terminologie, registres et dossiers partagés [334]. Le standard HL7 FHIR (*Fast Healthcare Interoperability Resources*, version R5 de 2023) fournit les ressources grâce auxquelles une

plateforme peut écrire dans le DSE des observations (pression artérielle, poids), des documents (la note de consultation) et des rendez-vous [305]. L'intégration est réalisable à grande échelle : à la Mayo Clinic, les consultations vidéo à domicile de 2020 ont fonctionné au moyen d'une application Zoom intégrée au dossier électronique Epic [342].

La figure 9.2 résume l'architecture. Les dispositifs à domicile transmettent, le plus souvent par Bluetooth, à une passerelle (*gateway* : le téléphone du patient ou un concentrateur), qui envoie les données à la plateforme par Internet. La plateforme applique les règles d'alerte et écrit les données dans le DSE, tandis que le clinicien consulte un tableau de bord, décide et répond au patient. La sécurité n'est pas un composant à part, mais une couche qui couvre toute la chaîne.

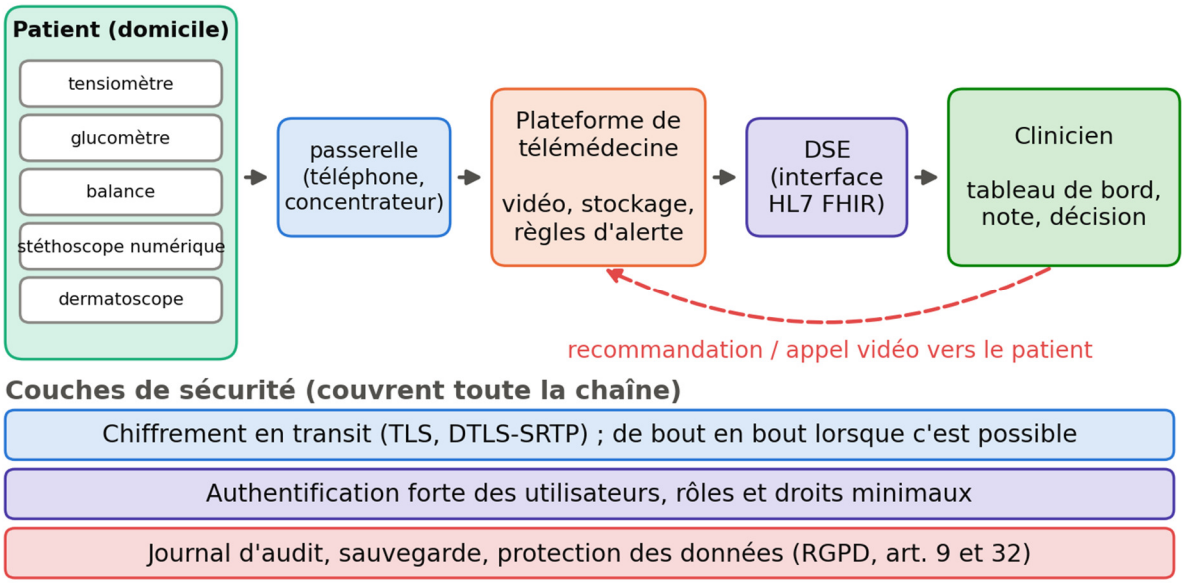


Figure 9.2. Architecture d'un système de téléconsultation et de télésurveillance, avec les couches de sécurité qui couvrent l'ensemble de la chaîne. Schéma didactique élaboré d'après [118,305,334].

9.3 Preuves d'efficacité 2020-2026

9.3.1 La pandémie de COVID-19 et l'adoption accélérée

La pandémie a changé l'échelle de la télémédecine au cours des premiers mois de 2020. Dans une base de données de 16,7 millions d'assurés aux États-Unis, 30,1 % de toutes les consultations ambulatoires de la période du 18 mars au 16 juin 2020 se sont déroulées par télémédecine, et le nombre hebdomadaire de téléconsultations a été multiplié par 23 par rapport à la période prépandémique [343]. Les écarts entre spécialités ont été importants : 68 % des endocrinologues ont utilisé la télémédecine, contre 9 % des ophtalmologistes. Selon l'affection, la part des téléconsultations a varié de 53 % pour la dépression à 3 % pour le glaucome [343]. L'écart a une explication clinique évidente : la dépression s'évalue par l'entretien, le glaucome par des mesures instrumentales.

En Europe, l'étude PRICOV-19 a inclus 5 065 cabinets de médecine de famille de 38 pays. Moins de la moitié d'entre eux (47,5 %) ont utilisé des consultations vidéo pendant la pandémie [344]. L'utilisation a été la plus élevée au Royaume-Uni, au Luxembourg, dans les pays scandinaves et en France (82,6-94,4 %) et la plus faible au Portugal, en Espagne, en Serbie, en Bosnie-Herzégovine, en Suisse et en Tchéquie (11,1-23,1 %). À l'échelle des pays, le seul facteur associé de façon statistiquement significative a été l'accès à Internet à un prix abordable [344].

Après la fin de l'état d'urgence, une partie des services s'est retirée. Une étude longitudinale portant sur 1 001 centres de santé mentale américains a constaté un recul de l'offre de télésanté de 81,6 % à 79,0 %, avec davantage de restrictions d'accès pour la thérapie et la gestion des médicaments [345].

La crise sanitaire a également réorganisé les flux d'information dans les spécialités chirurgicales. Une revue narrative de 23 études consacrées à la chirurgie programmée du cancer colorectal pendant la pandémie, dont l'auteur de ce manuel est coauteur, a rapporté des baisses du nombre d'interventions programmées allant de 14 % à 70 % selon les pays, des hospitalisations plus courtes et des délais d'attente plus longs avant l'admission. Dans un service de chirurgie de Timișoara, les interventions programmées pour cancer du côlon ont diminué de 42 % au cours de la première année de la pandémie [116]. Parmi les leçons formulées en vue d'une crise future figurent l'organisation en ligne des réunions multidisciplinaires et l'usage systématique de la télémédecine pour l'évaluation préopératoire et le suivi postopératoire [116]. L'étude ne mesure pas l'effet de la télémédecine sur les résultats cliniques. Son intérêt pour ce chapitre est autre : elle montre que, lorsque le contact physique devient risqué, la circulation de l'information (triage, scores de priorisation, réunions multidisciplinaires à distance) devient partie intégrante de l'acte médical.

9.3.2 Les preuves par spécialité

L'insuffisance cardiaque (IC) est le domaine où les preuves sont les plus nombreuses. Une méta-analyse de 92 études (65 sur la télésurveillance non invasive, 27 sur des dispositifs implantés), portant sur 36 549 patients suivis en moyenne 11,5 mois, a mis en évidence une réduction de 16 % de la mortalité toutes causes (OR 0,84, IC à 95 % 0,77-0,93) et de 19 % de la première hospitalisation pour IC (OR 0,81) [346]. Les auteurs notent toutefois que les méthodes de télésurveillance restent très diverses. Une seconde méta-analyse, limitée aux maladies cardiovasculaires, a trouvé pour l'association surveillance et consultation à distance dans l'IC une réduction de la mortalité cardiovasculaire (RR 0,83) et des hospitalisations d'origine cardiovasculaire (RR 0,71), sans effet sur les hospitalisations et la mortalité toutes causes. La consultation à distance seule n'a apporté aucun bénéfice [347]. La figure 9.3 met ces résultats côte à côte.

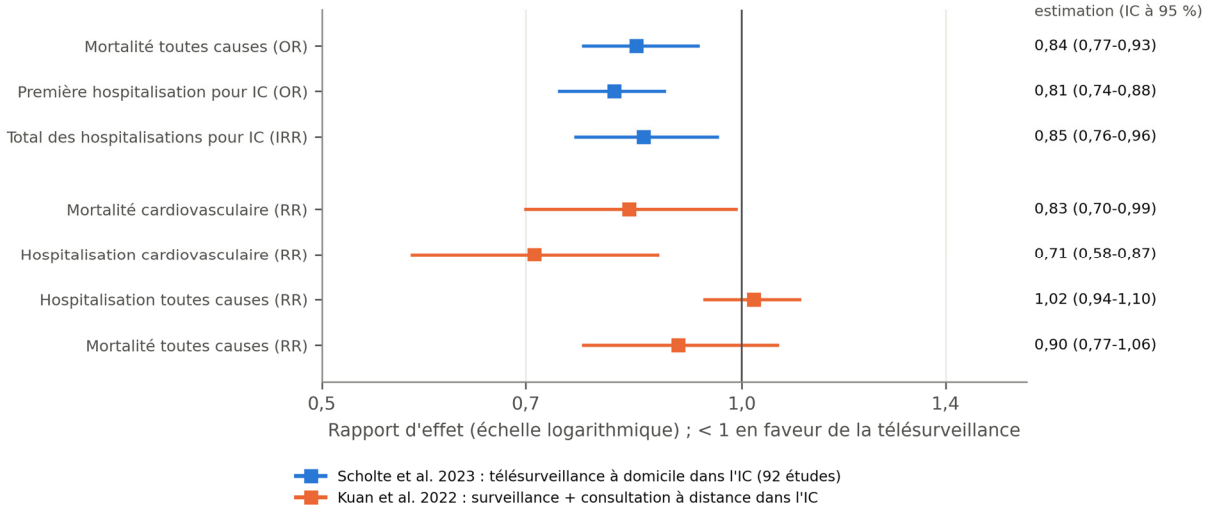


Figure 9.3. Effets cumulés de la télésurveillance dans l'insuffisance cardiaque, rapportés par deux méta-analyses. Les valeurs inférieures à 1 sont en faveur de la télésurveillance. OR, rapport de cotes ; IRR, rapport des taux d'incidence ; RR, risque relatif ; IC, intervalle de confiance. Données tirées de [346,347].

L'hypertension artérielle. Une méta-analyse de 23 études randomisées avec transmission automatique des valeurs tensionnelles mesurées à domicile a rapporté une baisse de la pression systolique mesurée au cabinet de 6,0 mmHg par rapport aux soins habituels [348]. Une méta-analyse en réseau de 24 ECR a estimé pour la télésurveillance une réduction de la pression systolique de 3,69 mmHg (IC à 95 % 1,57-5,82) par rapport aux soins habituels, mais n'a pas trouvé de différence par rapport à l'automesure à domicile sans transmission [349]. Ce résultat suggère qu'une partie du bénéfice vient de la mesure régulière, et pas seulement de la transmission.

Le diabète de type 2. Une revue de 246 études randomisées, dont 168 ont permis de calculer l'effet, a trouvé une baisse moyenne de l'HbA1c de 0,42 point de pourcentage, avec une hétérogénéité très importante (I^2 93 %). Les interventions qui comprenaient la télésurveillance ont eu des effets plus importants [350]. La téléconsultation synchrone, analysée séparément dans 27 études, a réduit l'HbA1c de 0,35 point de pourcentage, sans effet sur l'indice de masse corporelle, la pression artérielle ou le LDL-cholestérol [351].

La bronchopneumopathie chronique obstructive. La revue Cochrane de 2021 (29 études, 5 654 participants) est plus réservée : la surveillance à distance ajoutée aux soins habituels ne semble pas modifier le nombre d'exacerbations ni la qualité de vie, avec une certitude faible ou très faible, même si les réhospitalisations pour BPCO à 26 semaines ont probablement été réduites (HR 0,42, certitude modérée) [352]. La téléréadaptation respiratoire a en revanche donné des résultats semblables à ceux de la réadaptation en centre pour la distance parcourue en 6 minutes (différence moyenne 0,06 m), avec un taux d'achèvement de 93 %, contre 70 % pour les programmes classiques [353].

La psychiatrie. Une méta-analyse de 32 études randomisées (3 592 participants, 11 affections) n'a pas trouvé de différence globale entre la télépsychiatrie et la prise en charge en face à face. Par diagnostic, la télépsychiatrie a été supérieure dans les troubles dépressifs, et la consultation en face à face dans les troubles du comportement alimentaire [354].

La dermatologie. La méta-analyse citée plus haut a estimé une concordance diagnostique de 76 % entre la télédermatologie et la consultation directe pour l'ensemble des affections cutanées et de 73 % pour les cancers cutanés, sans différence entre le mode synchrone et le mode asynchrone [338].

L'accident vasculaire cérébral. Une méta-analyse portant sur 12 540 patients n'a pas trouvé de différence entre le télé-AVC (*telestroke*) et la prise en charge conventionnelle pour le délai porte-aiguille, la récupération fonctionnelle (score de Rankin modifié 0-2, OR 1,06), la mortalité à 90 jours ou l'hémorragie intracérébrale symptomatique [355]. Une étude allemande de non-infériorité de 2026, menée dans 15 hôpitaux de réseaux de télé-AVC, a comparé les visites neurologiques vidéo en service aux visites physiques chez les mêmes patients : le respect complet des indicateurs de qualité a été de 92 % en téléneurologie, contre 54 % lors de la visite au lit du malade [356]. L'étude n'était pas randomisée, et la comparaison porte sur la qualité de la documentation, non sur les résultats cliniques.

La radiologie. Dans une étude prospective portant sur 7 209 comptes rendus de TDM réalisés pendant les gardes, des divergences cliniquement importantes entre le compte rendu préliminaire et le compte rendu définitif sont apparues dans 6,1 % des comptes rendus de l'hôpital et dans 5,8 % de ceux de téléradiologie [357].

Tableau 9.2. Synthèse des preuves par spécialité (revues et études 2020-2026).

Domaine	Source	Type, taille	Résultat principal
Insuffisance cardiaque	[346]	méta-analyse, 92 études, 36 549 patients	mortalité OR 0,84
Insuffisance cardiaque	[347]	méta-analyse, 34 études dans l'analyse quantitative	mortalité CV RR 0,83, consultation seule sans effet
Hypertension	[348,349]	méta-analyses, 23 et 24 ECR	PA systolique -6,0 et -3,69 mmHg
Diabète de type 2	[350,351]	méta-analyses, 168 et 27 études	HbA1c -0,42 et -0,35 point de pourcentage
BPCO	[352]	Cochrane, 29 ECR	exacerbations inchangées, réhospitalisations HR 0,42
Psychiatrie	[354]	méta-analyse, 32 ECR	pas de différence globale
Dermatologie	[338]	méta-analyse, 155 études	concordance 76 %
AVC aigu	[355]	méta-analyse, 12 540 patients	résultats similaires
Radiologie	[357]	étude prospective, 7 209 comptes rendus	divergences 5,8 % contre 6,1 %

9.3.3 Ce qui fonctionne et ce qui ne fonctionne pas

Quelques régularités se dégagent de ces données. La télémédecine donne ses meilleurs résultats lorsque la transmission des données est reliée à une équipe qui réagit : dans l'IC, la surveillance suivie d'une consultation a réduit les événements cardiovasculaires, alors que la consultation seule ne l'a pas fait [347], et dans le diabète la composante de surveillance a amplifié l'effet [350]. Les effets sur les critères intermédiaires (pression artérielle, HbA1c) sont réels, mais modestes, de l'ordre de quelques mmHg ou de quelques dixièmes de point de pourcentage. L'hétérogénéité est importante [350], et les bénéfices dans l'IC proviennent surtout d'études à suivi court [347]. Enfin, l'équivalence avec la consultation en face à face est plus probable là où le diagnostic repose sur l'entretien ou sur des images (psychiatrie, dermatologie, radiologie) et moins probable là où il faut une palpation, une auscultation fine ou des mesures instrumentales.

9.4 Limites et risques

9.4.1 L'examen physique limité et les erreurs diagnostiques

La limite la plus directe est l'absence d'examen physique complet. À la Mayo Clinic, le diagnostic provisoire posé lors d'une consultation vidéo a été confirmé par l'examen en présentiel ultérieur dans 86,9 % des 2 393 nouveaux cas [342]. La concordance a varié de 77,3 % en oto-rhino-laryngologie à 96,0 % en psychiatrie et, selon les chapitres de la CIM-10, elle est descendue à 64,7 % pour les maladies de l'oreille [342]. Autrement dit, environ un cas sur huit a reçu un autre diagnostic lors de l'examen en présentiel. Le guide de l'OMS ajoute une limite plus difficile à mesurer : le canal vidéo réduit les signaux non verbaux, ce qui nuit à la confiance et à la vérification de la compréhension [334].

9.4.2 La continuité des soins

La même technologie peut produire des soins de qualité différente selon celui qui l'utilise. Dans une étude portant sur 27 686 enfants atteints d'infections respiratoires aiguës, des antibiotiques ont été prescrits dans 28,9 % des téléconsultations réalisées par le médecin de

famille habituel et dans 37,2 % de celles proposées par des entreprises de télémédecine directe au consommateur [358]. Les consultations répétées dans les 1 à 2 jours suivants ont elles aussi été plus rares lorsque la téléconsultation avait été assurée par le médecin qui connaissait le patient (5,0 % contre 8,0 %) [358]. La télémédecine intégrée à une relation médecin-patient existante n'est pas la même chose qu'une consultation ponctuelle avec un médecin inconnu.

9.4.3 La fracture numérique

Les patients dont les besoins de soins sont les plus grands sont souvent les moins préparés à la télémédecine. Une analyse d'un échantillon représentatif de 4 525 personnes âgées américaines a estimé que 38 % d'entre elles (environ 13 millions de personnes) n'étaient pas prêtes pour une consultation vidéo et 20 % pas même pour une consultation téléphonique, en raison de l'inexpérience technologique, de déficits auditifs et de communication ou d'une démence [359]. Chez les personnes de plus de 85 ans, la proportion de celles qui n'étaient pas prêtes pour la vidéo a atteint 72 %, et chez celles qui jugeaient leur santé mauvaise, 77 % [359]. L'utilisation de la télémédecine a été plus faible dans les communautés les plus touchées par la pauvreté [343], et une connexion Internet médiocre a limité l'accès aux consultations vidéo [336]. C'est pourquoi le guide de l'OMS demande plusieurs voies d'accès (plusieurs canaux de communication, des points d'accès tels que des bornes, le soutien des agents de santé communautaires) ainsi que l'évaluation de l'effet sur l'équité avant la mise en œuvre [334].

9.4.4 Confidentialité et responsabilité

Une revue de la littérature éthique et juridique de la période pandémique a montré que les problèmes les plus souvent discutés étaient le consentement éclairé et l'autonomie (87 % des articles), la vie privée (78 %), la protection et la sécurité des données (74 %) et la responsabilité professionnelle (70 %) [360]. Les auteurs concluent que les services de télémédecine peuvent être utilisés, pour l'instant, comme des outils complémentaires des soins traditionnels. Le guide de l'OMS recommande de clarifier, avant le lancement, le cadre d'autorisation d'exercice des professionnels, la responsabilité médicale et les modalités d'audit des services [334]. Des risques pour la confidentialité apparaissent aussi en dehors de la plateforme : le patient qui parle depuis un bureau ouvert, un membre de la famille qui écoute, un écran partagé alors qu'une autre fenêtre est ouverte.

9.4.5 Les fraudes

En juillet 2022, le Bureau de l'inspecteur général du ministère de la Santé des États-Unis a émis une alerte spéciale concernant les pratiques frauduleuses de certaines entreprises de télémédecine, dans le contexte de plusieurs affaires de fraude portant sur des milliards de dollars et poursuivies par les autorités [361]. Une analyse des études fédérales américaines de ces dernières années a toutefois conclu que peu de données indiquent que la télésanté serait associée à des niveaux de fraude et d'abus plus élevés que les soins classiques [362]. Pour l'étudiant, la leçon est pratique : la note de téléconsultation doit refléter exactement ce qui a été fait, combien de temps cela a duré et par quel canal, car la facturation et l'audit reposent sur elle.

9.5 Le cadre juridique et le remboursement

9.5.1 La Roumanie : la situation au 27 septembre 2026

La télémédecine est entrée dans la législation roumaine par l'ordonnance d'urgence du Gouvernement (Ordonanța de urgență a Guvernului, OUG) n° 196/2020 (18 novembre 2020, Monitorul Oficial, Journal officiel de la Roumanie, n° 1108 du 19 novembre 2020), qui a introduit dans la loi n° 95/2006 sur la réforme dans le domaine de la santé (Legea nr. 95/2006 privind reforma în domeniul sănătății) les articles 30^1 à 30^11 [363]. La loi définit la télémédecine comme l'ensemble des services médicaux fournis à distance, sans présence physique simultanée du personnel médical et du patient, à des fins de diagnostic, de traitement, de surveillance ou de prévention, au moyen des technologies de l'information et des communications électroniques, dans des conditions de sécurité. Les formes réglementées sont la téléconsultation, la téléexpertise, la téléassistance médicale, la téléradiologie, la télépathologie et la télésurveillance [363,364].

Les normes méthodologiques d'application ont été approuvées par la décision du Gouvernement (Hotărârea Guvernului, HG) n° 1133 du 14 septembre 2022 (Monitorul Oficial n° 909 du 15 septembre 2022) [364]. Plusieurs dispositions de l'annexe 2 de ces normes ont un effet direct sur la pratique :

- Hormis la médecine d'urgence, l'accès à la plateforme ou la fourniture du service par télémédecine exigent le consentement exprès du patient, prouvé par tout moyen de preuve, et le patient peut refuser le service (art. 4).
- La télémédecine n'est pas utilisée lorsque le médecin et le patient ne parlent pas suffisamment bien la même langue (art. 9, al. 1).
- Les services sont programmés comme ceux en face à face, de sorte que le médecin ait accès au dossier du patient au moment de la consultation (art. 9, al. 3).
- Le compte rendu du service est consigné dans le registre des consultations ou dans le dossier du patient, avec le moyen de communication utilisé, la date et la plage horaire, les prescriptions et les documents délivrés, l'identité du personnel médical participant et les éventuels incidents techniques (art. 10).

Un détail terminologique mérite d'être retenu. Dans les normes, la téléassistance désigne l'intégration en temps réel de données ou d'avis de professionnels se trouvant à distance en vue de la réalisation d'actes médicaux [364]. Le sens juridique roumain est donc proche de l'assistance d'une procédure par un expert, et non du bouton d'alarme de l'aide sociale décrit à la section 9.1.

Le remboursement par le Fondul național unic de asigurări sociale de sănătate (Fonds national unique d'assurance sociale de santé) suit le contrat-cadre (Contractul-cadru) approuvé par la décision du Gouvernement n° 521/2023 et ses modalités d'application, l'arrêté du ministère de la Santé et de la CNAS (Ordinul MS/CNAS) n° 1857/441/2023 [365,366]. Dans la version consolidée des modalités d'application publiée par la CNAS en août 2025, le médecin de famille peut assurer à distance, "par tout moyen de communication", des consultations pour les maladies à potentiel endémo-épidémique qui nécessitent un isolement et des consultations pour les affections chroniques, dans les mêmes conditions qu'au cabinet, en les consignant dans le registre et dans le dossier avec le moyen de communication et la plage horaire [366]. En consultation externe de spécialité, les consultations à distance pour affections chroniques sont assurées et remboursées comme celles du cabinet, sans lettre d'orientation (bilet de trimitere), et les thérapies psychiatriques et le conseil génétique peuvent également être assurés à distance [366,367]. Les

modifications apportées aux modalités d'application après août 2025 n'ont pas été vérifiées pour ce chapitre, et le lecteur doit consulter la version en vigueur à la date de la consultation.

La situation en Roumanie en 2026

- Le financement de la numérisation de la santé au titre du Planul Național de Redresare și Reziliență (plan national pour la reprise et la résilience, PNRR) relève de la composante 7, Transformare digitală (transformation numérique), investissement I3, "Realizarea sistemului de eHealth și telemedicină" (mise en place du système de e-santé et de télémédecine), et non de la composante consacrée à la santé [368].
- Le sous-investissement I3.3 a financé des systèmes informatiques et une infrastructure numérique pour 200 établissements de santé publics, avec une enveloppe de 493 900 000 lei hors TVA [326].
- Le projet de modernisation de la plateforme informatique de la CNAS (I3.1) a été déclaré achevé le 31 août 2026 [323].
- Au 27 septembre 2026, aucune source officielle vérifiée n'était disponible sur les cibles physiques atteintes en télémédecine grâce au PNRR.

9.5.2 Repères européens

L'organisation des services médicaux reste de la compétence des États membres, mais plusieurs actes de l'Union européenne s'appliquent directement à la télémédecine. La directive 2011/24/UE relative à l'application des droits des patients en matière de soins de santé transfrontaliers, en vigueur, dispose à son art. 3, point d), que, dans le cas de la télémédecine, les soins de santé sont considérés comme dispensés dans l'État membre où le prestataire de soins de santé est établi [369]. Cette règle a des conséquences sur la loi applicable et sur la responsabilité. La directive permet également à l'État membre d'affiliation d'appliquer aux soins reçus par télémédecine les mêmes conditions de remboursement que dans son propre système (art. 7, par. 7) [369].

Le règlement (UE) 2025/327 relatif à l'espace européen des données de santé (EHDS), publié en mars 2025, reconnaît la télémédecine comme un outil de plus en plus important et organise l'échange transfrontalier des données de santé au moyen de l'infrastructure MyHealth@EU [23]. Le RGPD reste applicable à toutes les formes de télémédecine [118].

9.6 La documentation de la téléconsultation

Une téléconsultation non documentée est, sur le plan médical et juridique, une téléconsultation qui n'a pas eu lieu. En Roumanie, le contenu minimal de la consignation est fixé par l'art. 10 de l'annexe 2 des normes relatives à la télémédecine (section 9.5) [364]. Le guide de l'OMS exige des mécanismes d'identification réciproque du médecin et du patient et de traçabilité des échanges et des décisions antérieurs [334]. Le même guide considère que, dans la plupart des pays, le consentement est implicite lorsque le patient sollicite le service et que seule la téléconsultation initiée par le professionnel exige un consentement documenté, avec la communication explicite des limites de la télémédecine [334]. Les normes roumaines sont plus strictes : hors urgence, le consentement doit être exprès et prouvé, quel que soit l'initiateur de la consultation [364]. Les lignes directrices australiennes de pratique en télédermatologie, mises à jour en 2025, comportent des chapitres distincts sur la sélection des patients et le consentement, les standards des images, la confidentialité, la documentation et la conservation des images [370].

Un point faible fréquent est la consignation des conseils de sécurité (*safety-netting*), c'est-à-dire des instructions sur les signes qui doivent amener le patient à revenir ou à appeler le 112. Dans une analyse de 295 consultations enregistrées en médecine de famille, les conseils de sécurité ont été donnés oralement dans deux tiers des consultations, mais n'ont été documentés que dans un tiers d'entre elles, couvrant 20,3 % des problèmes évalués [371]. En télémédecine, où l'examen est incomplet, ces instructions importent davantage qu'au cabinet.

Une note de téléconsultation contient au minimum :

1. **l'identification** : le patient (nom, date de naissance ou autre identifiant vérifié), le médecin (nom, spécialité, code de cachet du médecin, ou parafă), la date, l'heure de début et de fin.
2. **la localisation** du patient au moment de la consultation et un numéro de téléphone pour le rappeler.
3. **la modalité** : vidéo, téléphone, message asynchrone, plateforme utilisée, personnes présentes.
4. **le consentement** à la consultation à distance et, séparément, à l'enregistrement, s'il y en a un.
5. **les données cliniques** : le motif, l'anamnèse, ce qui a pu être observé ou mesuré (y compris les valeurs transmises par des dispositifs).
6. **les limites** : ce qui n'a pas pu être examiné et l'incidence sur la conclusion.
7. **l'évaluation et le plan** : diagnostic présumé, recommandations, prescriptions, examens complémentaires, orientations.
8. **la sécurité** : les signes d'alarme expliqués et le délai de réévaluation ou d'examen en présentiel.
9. **la signature** du médecin (électronique, là où le système le permet).

Exemple

Note de téléconsultation (patient fictif, données inventées). Date : 14.10.2026, 10 h 05-10 h 22. Médecin : Dr A.B., médecine interne, parafă [fictive]. Patient : I.P., 67 ans, identité vérifiée par la date de naissance et le CNP (code numérique personnel) figurant au dossier. Localisation : domicile, département de Timiș, téléphone de secours confirmé. Modalité : vidéo sur la plateforme de l'établissement, sans enregistrement, patient seul dans la pièce. Consentement exprès à la consultation à distance, exprimé oralement et consigné. Incidents techniques : aucun. Motif : contrôle de l'hypertension. Valeurs transmises au cours des 7 derniers jours : moyenne 148/88 mmHg. Ne se plaint ni de céphalées, ni de dyspnée, ni de douleur thoracique. Observé en vidéo : parole fluide, pas d'œdème visible aux chevilles (examen limité). Limites : l'auscultation et la mesure de la pression artérielle par le médecin n'ont pas pu être réalisées. Plan : augmentation de la dose de l'antihypertenseur selon le schéma, poursuite de l'automesure deux fois par jour. Ordonnance électronique émise. Sécurité : consigne donnée au patient d'appeler le 112 en cas de douleur thoracique, de faiblesse d'un membre ou de troubles de la parole. Réévaluation en vidéo dans 2 semaines, examen en présentiel avec ECG dans un mois. Signature électronique : Dr A.B.

Pièges fréquents

- Considérer toute application qui affiche des données de santé comme de la télémédecine. Sans un professionnel qui reçoit et utilise les données, il s'agit d'autosurveillance.
- Confondre le chiffrement en transit et le chiffrement de bout en bout. Certaines plateformes, par exemple Jitsi Meet sans E2EE activé, déchiffrent le flux sur le serveur.
- Lire une méta-analyse favorable comme une garantie valable pour n'importe quel produit. Les effets dépendent de la composante de surveillance, de l'équipe qui réagit et de la population.
- Omettre dans la note les limites de l'examen et les conseils de sécurité.
- Supposer que tous les patients peuvent utiliser la vidéo. Les personnes âgées, les personnes malentendantes ou sans connexion Internet stable ont besoin de solutions de remplacement.
- Poursuivre la téléconsultation lorsque des signes d'alarme apparaissent, au lieu d'orienter immédiatement le patient vers les urgences.
- Employer le terme de téléassistance au sens de l'aide sociale lorsqu'on discute de la législation roumaine, où la téléassistance médicale désigne le soutien d'un expert, en temps réel, pour la réalisation d'une procédure.

À retenir

- La télémédecine est la fourniture de services médicaux à distance au moyen des TIC, en mode synchrone, asynchrone ou par surveillance.
- La classification de l'OMS de 2023 code la télémédecine dans le sous-groupe 2.4, qui compte quatre interventions.
- Les bénéfices les plus nets apparaissent dans la télésurveillance de l'insuffisance cardiaque, lorsque les données parviennent à une équipe qui réagit.
- Dans l'hypertension et le diabète, les effets sont réels, mais modestes.
- En psychiatrie, en dermatologie et en radiologie, les résultats sont proches de ceux obtenus en face à face.
- Environ un diagnostic sur huit posé par vidéo a changé lors de l'examen en présentiel dans une cohorte de plus de 2 000 patients.
- La note de téléconsultation consigne la modalité, le consentement, les limites et les signes d'alarme expliqués.
- En Roumanie, la télémédecine est réglementée par l'OUG n° 196/2020 et par les normes approuvées par la HG n° 1133/2022, qui exigent un consentement exprès hors urgence.

Questions d'autoévaluation

1. (QCS) Un médecin de famille photographie au dermatoscope une lésion pigmentée et envoie l'image à un dermatologue, qui répond le lendemain. La forme de télémédecine est :

a) une téléconsultation synchrone
b) une téléexpertise asynchrone
c) une télésurveillance
d) une téléassistance
e) une autosurveillance

2. (QCS) Dans la classification de l'OMS de 2023, la surveillance à distance des données du patient par le prestataire de soins a le code :

a) 1.4.2
b) 2.4.1
c) 2.4.2
d) 2.4.4
e) 4.1.2

3. (QCS) Dans la méta-analyse de Scholte et collaborateurs (2023), la télésurveillance à domicile dans l'insuffisance cardiaque a été associée à :

a) une augmentation de la mortalité toutes causes
b) une réduction de la mortalité toutes causes d'environ 16 %
c) aucune différence pour la première hospitalisation pour insuffisance cardiaque
d) des effets uniquement pour les dispositifs implantés
e) une réduction de la mortalité de 50 %

4. (QCM) Lesquels des éléments suivants doivent être consignés dans une note de téléconsultation ?

a) la modalité (vidéo, téléphone) et la plateforme utilisée
b) les limites de l'examen à distance
c) les conseils de sécurité et les signes d'alarme expliqués
d) le mot de passe du compte du patient sur la plateforme
e) le consentement à la consultation à distance

5. (QCM) Lesquelles des affirmations sur les limites de la télémédecine sont étayées par les données présentées dans le chapitre ?

a) le diagnostic provisoire posé par vidéo a été confirmé lors de l'examen en présentiel dans environ 87 % des cas, dans une cohorte de la Mayo Clinic
b) environ 38 % des personnes âgées américaines n'étaient pas prêtes pour une consultation vidéo
c) le chiffrement en transit garantit que le serveur ne peut pas accéder au flux
d) les téléconsultations pédiatriques proposées par des entreprises directes au consommateur ont été associées à davantage de prescriptions d'antibiotiques que celles réalisées par le médecin de famille
e) il est prouvé que les fraudes sont plus fréquentes en télémédecine que dans les soins classiques

CHAPITRE 10

Chapitre 10. La santé mobile, les dispositifs portables et les modèles de langage en consultation

Objectifs

À la fin de ce chapitre, l'étudiant sera capable :

- de distinguer une application de bien-être d'une application qui est un dispositif médical, à partir de la destination déclarée par le fabricant
- d'interpréter les résultats d'une méta-analyse sur l'efficacité des applications mobiles (différence moyenne standardisée, NNT, certitude des preuves)
- d'énumérer les capteurs d'un dispositif portable et d'apprécier, pour chacun, le niveau d'exactitude documenté
- d'appliquer les étapes du guide MDCG 2019-11 et la règle 11 du MDR pour déterminer la classe de risque d'une application
- de décrire le trajet d'une mesure, du capteur au dossier électronique, et les normes et standards qui le rendent possible (ISO/IEEE 11073, Bluetooth, FHIR)
- d'utiliser de manière responsable un grand modèle de langage : vérifier les affirmations et les références, protéger les données du patient et déclarer l'utilisation

Le téléphone que le patient a dans sa poche est devenu, sans que personne l'ait planifié, l'instrument de mesure le plus répandu en médecine. Il compte les pas, reçoit le pouls transmis par la montre, affiche la glycémie mesurée par le capteur posé sur le bras et, depuis la fin de 2022, peut tenir une conversation fluide sur des symptômes. Le chapitre suit cette succession : l'application, le capteur, le cadre réglementaire, les normes et standards de transfert et, pour finir, les modèles de langage utilisés en consultation. Les méthodes de l'intelligence artificielle proprement dites (apprentissage automatique, évaluation, explicabilité, AI Act) sont traitées au chapitre 13. La télésurveillance organisée par un service médical a été décrite au chapitre 9.

10.1 La santé mobile

10.1.1 Définitions et catégories d'applications

La stratégie mondiale de l'Organisation mondiale de la santé (OMS) définit la santé numérique (*digital health*) comme le domaine de connaissances et de pratique lié au développement et à l'utilisation des technologies numériques pour améliorer la santé. Le document précise qu'elle étend le concept d'e-santé aux consommateurs numériques et à une large gamme d'appareils intelligents et connectés [36]. La santé mobile (*mobile health*, mHealth) est la partie de ce domaine qui utilise les téléphones, les tablettes et les dispositifs portables.

Pour le médecin, la classification la plus utile n'est pas la classification technique, mais celle qui repose sur la destination déclarée (*intended purpose*). Le guide européen MDCG 2019-11, révisé en juin 2025, indique que les logiciels destinés exclusivement à des fins non médicales, y

compris les applications de bien-être ou de fitness, ne sont pas qualifiés de logiciels dispositifs médicaux [372]. Le même guide donne des exemples de logiciels qui le sont : une application pour montre connectée qui envoie des alertes lorsqu'elle reconnaît des battements cardiaques irréguliers, un logiciel qui calcule la dose d'insuline et commande la pompe, une application qui contribue au traitement de la schizophrénie par la surveillance des symptômes et des interventions personnalisées [372]. Pour les applications sans finalité médicale, le règlement relatif à l'espace européen des données de santé (EHDS) a introduit une définition propre, celle d'application de bien-être (*wellness application*), c'est-à-dire un logiciel, seul ou combiné à du matériel, destiné par le fabricant à être utilisé par une personne physique pour traiter des données de santé électroniques à des fins autres que la prestation de soins de santé [23].

Il en résulte trois catégories pratiques. La première comprend les applications de mode de vie (pas, sommeil, alimentation, méditation), sans finalité médicale déclarée. La deuxième réunit les applications qui sont des dispositifs médicaux, porteuses du marquage CE, depuis les carnets de glycémie avec calcul de dose jusqu'aux thérapies numériques prescrites. La troisième, plus restreinte, inclut les dispositifs médicaux numériques remboursés par un système public d'assurance, dont l'exemple le plus étudié est le modèle allemand des DiGA (section 10.3.3).

10.1.2 Ce que disent les preuves 2020-2026

Le nombre d'essais randomisés a tellement augmenté que la synthèse se fait désormais au niveau des revues de revues (*umbrella review*). La plus vaste, publiée en 2026, a inclus 78 revues systématiques avec méta-analyse, couvrant 496 études primaires et 177 373 participants adultes atteints de maladies non transmissibles ou à risque de les développer [373]. Par rapport aux groupes témoins, les applications ont réduit la pression artérielle systolique (différence moyenne standardisée, *standardized mean difference*, SMD -0,444) et diastolique (SMD -0,414), l'hémoglobine glyquée (SMD -0,587), le poids et le temps sédentaire. Les effets sur l'anxiété (SMD -0,215) et la dépression (SMD -0,097) ont été faibles. Aucun effet n'a été trouvé sur la mortalité cardiovasculaire ni sur les lipides sanguins, et l'hétérogénéité a été très importante (I^2 de 91-96 % pour la pression artérielle et les indicateurs glycémiques) [373].

Une seconde synthèse, portant sur 34 méta-analyses et 235 essais randomisés menés dans 52 pays, a appliqué des critères plus stricts de solidité des preuves. Le seul effet "convaincant" a été la réduction de l'hémoglobine glyquée dans le diabète de type 2 (d = 0,44), et 42 % des effets analysés n'étaient pas statistiquement significatifs. Les auteurs ont également relevé que les effets indésirables sont rarement rapportés [374].

Par maladie, le tableau est le suivant :

- **Le diabète de type 2.** Dans une méta-analyse d'essais randomisés, les interventions par application ont réduit l'HbA1c de 0,42 point de pourcentage par rapport au groupe témoin, celles par SMS de 0,37, tandis que celles par site web n'ont pas eu d'effet statistiquement significatif (-0,09) [375].
- **L'hypertension artérielle.** Une revue de 74 études et 92 686 participants a trouvé des baisses de la pression systolique et diastolique pour les interventions mHealth, et l'observance médicamenteuse ne s'est améliorée que dans le groupe mHealth [376].
- **Le sevrage tabagique.** Une méta-analyse de 2026 (31 études, 12 802 participants) a estimé que l'application utilisée seule pourrait augmenter l'abstinence continue à 6 mois par rapport à l'absence de soutien ou à un soutien minimal (RR 2,85), mais la certitude des preuves était faible et la comparaison ne reposait que sur 4 études [377].

- **La dépression et l'anxiété.** Dans 176 essais randomisés, les applications ont eu des effets faibles sur les symptômes dépressifs (g = 0,28, NNT 11,5) et d'anxiété généralisée (g = 0,26, NNT 12,4). Les effets étaient plus importants lorsque l'application comportait des éléments de thérapie cognitivo-comportementale [378].

Que signifient ces chiffres pour un étudiant ? Un NNT de 11,5 indique qu'en moyenne 11 à 12 personnes doivent utiliser l'application pour qu'une personne de plus ait une réponse clinique par rapport au groupe témoin. L'effet est réel, mais modeste, et il a été obtenu dans les conditions contrôlées d'un essai.

10.1.3 La qualité des applications et la rétention des utilisateurs

Dans les magasins d'applications, la qualité ne se déduit pas du nombre d'étoiles. L'instrument le plus utilisé pour l'évaluer est l'échelle MARS (*Mobile App Rating Scale*), qui note quatre dimensions : l'engagement de l'utilisateur, la fonctionnalité, l'esthétique et la qualité de l'information. Une étude de validation portant sur 1 299 applications de 15 domaines de santé a confirmé la structure de l'échelle, avec une fiabilité bonne ou excellente (coefficient oméga 0,79-0,93) et un accord élevé entre évaluateurs (CCI 0,82) [379]. Une revue parapluie de 2024 a constaté que MARS était le cadre le plus souvent utilisé, mais a observé que la plupart des cadres ne couvrent pas tous les critères identifiés. La protection des données et la sécurité ont été le critère le plus souvent mentionné, dans 13 revues sur 15 [380].

Le second problème est l'abandon. Une revue systématique de 99 études sur les applications destinées aux maladies non transmissibles a calculé un score moyen d'observance de 56,0 %, rapporté à l'utilisation prévue par le fabricant [381]. Les facteurs associés à une meilleure observance ont été la personnalisation du contenu, les notifications individualisées, une conception stable et facile à utiliser et un soutien humain complémentaire. Une méta-analyse de 2026 des essais randomisés portant sur des applications pour la dépression et l'anxiété a trouvé une activation initiale de 92,4 %, mais une observance de seulement 61,8 % et un taux de perdus de vue de 28,4 % lors de l'évaluation de suivi. Les pertes étaient plus faibles lorsqu'il existait des rappels et un contact humain [382]. Ces chiffres proviennent d'études dans lesquelles les participants sont sélectionnés, motivés et suivis activement. C'est pourquoi ils ne doivent pas être lus comme des estimations de l'utilisation réelle.

10.2 Dispositifs portables et capteurs

10.2.1 Ce que mesure un dispositif portable

Un dispositif portable (*wearable*) est un capteur porté sur le corps, en général au poignet, au doigt, sur la poitrine ou sur le bras, qui transmet ses données à un téléphone. Le trajet complet d'une mesure, du capteur au médecin, est présenté dans la figure 10.1. Les points signalés sur le schéma (le consentement, le chiffrement, le contrôle d'accès) sont les endroits où se décide si les données peuvent être utilisées cliniquement.

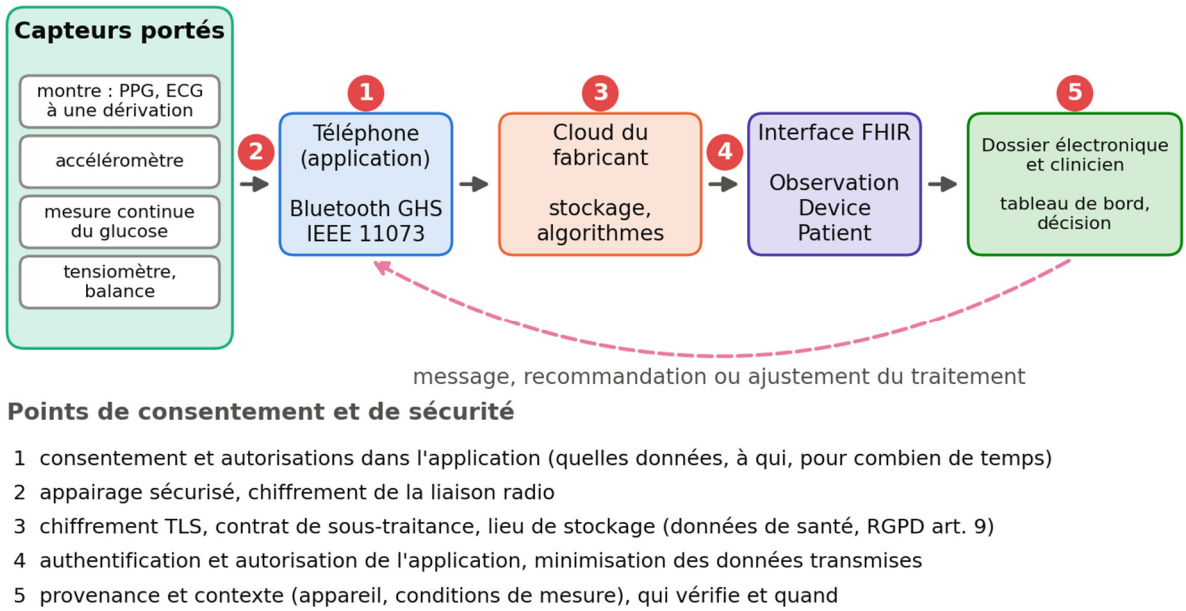


Figure 10.1. Flux des données générées par le patient : capteur, téléphone, plateforme du fabricant, dossier électronique et clinicien, avec les points de consentement et de sécurité. Schéma didactique élaboré d'après [118,383-385].

Les principaux capteurs sont les suivants :

- **L'accéléromètre** (souvent associé au gyroscope) enregistre le mouvement selon trois axes. On en dérive le nombre de pas, l'intensité de l'activité, les chutes et, indirectement, le sommeil.
- **La photopléthysmographie** (*photoplethysmography*, PPG) émet une lumière verte ou rouge dans la peau et mesure la variation de la lumière réfléchie à chaque onde de pouls. À partir du signal, on calcule la fréquence cardiaque, l'irrégularité du rythme et, avec différentes longueurs d'onde, la saturation en oxygène (SpO2).
- **L'électrocardiogramme à une dérivation** s'obtient lorsque l'utilisateur touche avec un doigt l'électrode de la montre et ferme ainsi le circuit entre les deux bras. Le résultat est un tracé court, semblable à la dérivation I.
- **Le capteur de température** mesure la température cutanée, non la température centrale. Son utilité tient surtout aux tendances.
- **Le système de mesure continue du glucose** (*continuous glucose monitoring*, CGM) comporte un filament introduit dans le tissu sous-cutané et mesure le glucose interstitiel à intervalles de quelques minutes. De par sa destination déclarée, c'est un dispositif médical.

Tableau 10.1. Capteurs des dispositifs portables, exactitude documentée et utilisations cliniques étayées par des preuves. Source : [238,242,243,386-388].

Capteur	Ce qu'il mesure	Exactitude rapportée	Utilisation étayée par des preuves
Accéléromètre	pas, activité, sommeil	pas sous-estimés (erreur de -9 % à 12 %), sommeil surestimé	récupération après chirurgie, activité physique
PPG	pouls, rythme, SpO2	pouls avec une erreur moyenne de ±3 %	dépistage de la FA (sensibilité 95 %, spécificité 97 %)
ECG à une dérivation	tracé semblable à la dérivation I	pour la FA, comparable à la PPG	vérification d'un rythme irrégulier
Température cutanée	température de la peau, tendances	ce n'est pas la température centrale	récupération postopératoire, en association
SpO2 optique	saturation en oxygène	différence moyenne allant jusqu'à 2 %	indicative, non diagnostique
Pression sans brassard	PA estimée à partir de la PPG	validation non standardisée	pas encore recommandée en clinique
CGM	glucose interstitiel	dispositif médical, efficacité dans des ECR	HbA1c réduite dans le diabète de type 1 et de type 2

10.2.2 Exactitude et validation

Une revue parapluie "vivante" (mise à jour en continu) a rassemblé 24 revues systématiques et 249 études de validation portant sur plus de 430 000 participants [386]. Seuls 11 % environ des dispositifs commercialisés avaient été validés pour au moins une mesure. Pour la fréquence cardiaque, l'erreur moyenne a été de ±3 %, et pour la détection des arythmies, la sensibilité et la spécificité cumulées ont été respectivement de 100 % et de 95 %. Les pas ont été en général sous-estimés, tout comme la dépense énergétique, alors que la consommation maximale d'oxygène a été surestimée, de même que la durée totale du sommeil, avec des erreurs typiquement supérieures à 10 %. Pour la SpO2, la différence moyenne absolue a atteint 2,0 % [386].

La couleur de la peau compte pour les capteurs optiques. Une méta-analyse de 23 études d'oxymétrie de pouls (197 353 paires SpO2-sang artériel) a montré que l'erreur quadratique moyenne dépassait le seuil réglementaire de 3 % dans tous les groupes de pigmentation (3,96 %, 4,71 % et 4,15 % pour les peaux claires, moyennes et foncées) [242]. Pour le pouls mesuré par les dispositifs portables, la même analyse n'a pas trouvé d'erreur systématique liée à la pigmentation, bien que les limites d'agrément aient été larges.

Pour la pression artérielle mesurée sans brassard, les preuves sont plus faibles. Dans une revue de 16 études et 15 dispositifs, seuls 8 dispositifs ont présenté une erreur moyenne inférieure à 5 mmHg. La validation selon des protocoles standard et en vie réelle reste la condition de leur utilisation clinique [387].

À l'opposé se trouve le CGM, évalué dans des dizaines d'essais randomisés. Une méta-analyse de 64 essais randomisés a trouvé pour le CGM en temps réel, par rapport à l'autosurveillance par glucomètre, une réduction de l'HbA1c de 0,24 point de pourcentage dans le diabète de type 1 et de 0,40 dans le diabète de type 2. Le temps passé dans la plage cible (70-180 mg/dl) a augmenté d'environ 5,5 points de pourcentage dans les deux types de diabète [388].

10.2.3 Utilisations cliniques : arythmies et surveillance postopératoire

La fibrillation atriale (FA) est l'application la mieux documentée. Une méta-analyse de 2025, portant sur 26 études et 17 349 patients, a rapporté pour les montres connectées une sensibilité cumulée de 95 % et une spécificité de 97 %, avec des performances comparables entre les modèles fondés sur la PPG et ceux dotés d'un ECG [238]. L'étude Fitbit Heart, un essai clinique mené entièrement à distance, a recruté 455 699 adultes. La notification de rythme irrégulier est apparue chez 1 % des participants et chez 4 % de ceux âgés de plus de 65 ans. Parmi ceux qui ont ensuite porté un patch ECG analysable, 32,2 % présentaient une FA, et la valeur prédictive positive d'une notification survenue au cours même de la surveillance par patch a été de 98,2 % [235]. Les chiffres montrent aussi la limite : chez environ deux tiers des personnes ayant reçu une notification qui ont porté le patch, celui-ci n'a pas enregistré de FA pendant la semaine de surveillance. Le diagnostic se confirme sur l'ECG, et la décision d'anticoagulation revient au médecin.

La seconde utilisation, plus récente, est le suivi de la récupération après chirurgie. Une revue systématique de l'auteur de cet ouvrage a analysé la surveillance continue par capteurs portables après chirurgie colorectale [243]. Neuf études portant sur 778 patients ont été incluses : un essai randomisé, sept cohortes prospectives et une analyse rétrospective. Cinq études ont utilisé uniquement le nombre de pas, et quatre l'ont associé à la fréquence cardiaque ou à la température cutanée, sur une durée médiane de port de 6 jours. Au moins 1 000 pas le premier jour postopératoire ont été associés à une durée d'hospitalisation plus courte (OR 0,63), et un protocole ERAS (*Enhanced Recovery After Surgery*) complété par des bracelets connectés a raccourci de 1,1 jour l'hospitalisation définie par le protocole. L'inactivité préopératoire (moins de 5 000 pas par jour) et le retour incomplet au niveau d'activité initial ont prédit les complications et la réhospitalisation à 30 jours. La revue a conclu que ces données sont exploitables et cliniquement utiles, mais qu'il manque des définitions standardisées des critères de jugement, des algorithmes ouverts de traitement du signal et une validation multicentrique [243].

10.3 Le logiciel dispositif médical

10.3.1 La qualification : quand une application devient-elle un dispositif médical ?

Le règlement (UE) 2017/745 relatif aux dispositifs médicaux (MDR) s'applique également aux logiciels [284]. La question pratique, "mon application est-elle un dispositif médical ?", trouve sa réponse dans le guide MDCG 2019-11 rev.1 [372]. Le critère central est la destination médicale déclarée par le fabricant. Le risque de nuire n'est pas, en soi, un critère de qualification. Le guide propose cinq étapes de décision :

1. Le produit est-il un logiciel au sens du guide ?
2. S'agit-il d'un dispositif de l'annexe XVI du MDR, d'un accessoire ou d'un logiciel qui commande un dispositif médical physique ou influence son utilisation ? Si oui, il suit le régime de ce dispositif.
3. Réalise-t-il une action sur les données, au-delà du stockage, de l'archivage, de la communication, de la simple recherche ou de la compression sans perte ?
4. Cette action est-elle au bénéfice d'un patient individuel ? Les logiciels qui se contentent d'agréger des données populationnelles, de reproduire des recommandations générales ou d'alimenter des registres épidémiologiques ne remplissent pas cette condition.
5. Le logiciel entre-t-il dans la définition du dispositif médical (diagnostic, prévention, contrôle, prédiction, pronostic, traitement ou atténuation d'une maladie) ?

La figure 10.2 transpose ces étapes et la règle 11 en un arbre de décision.

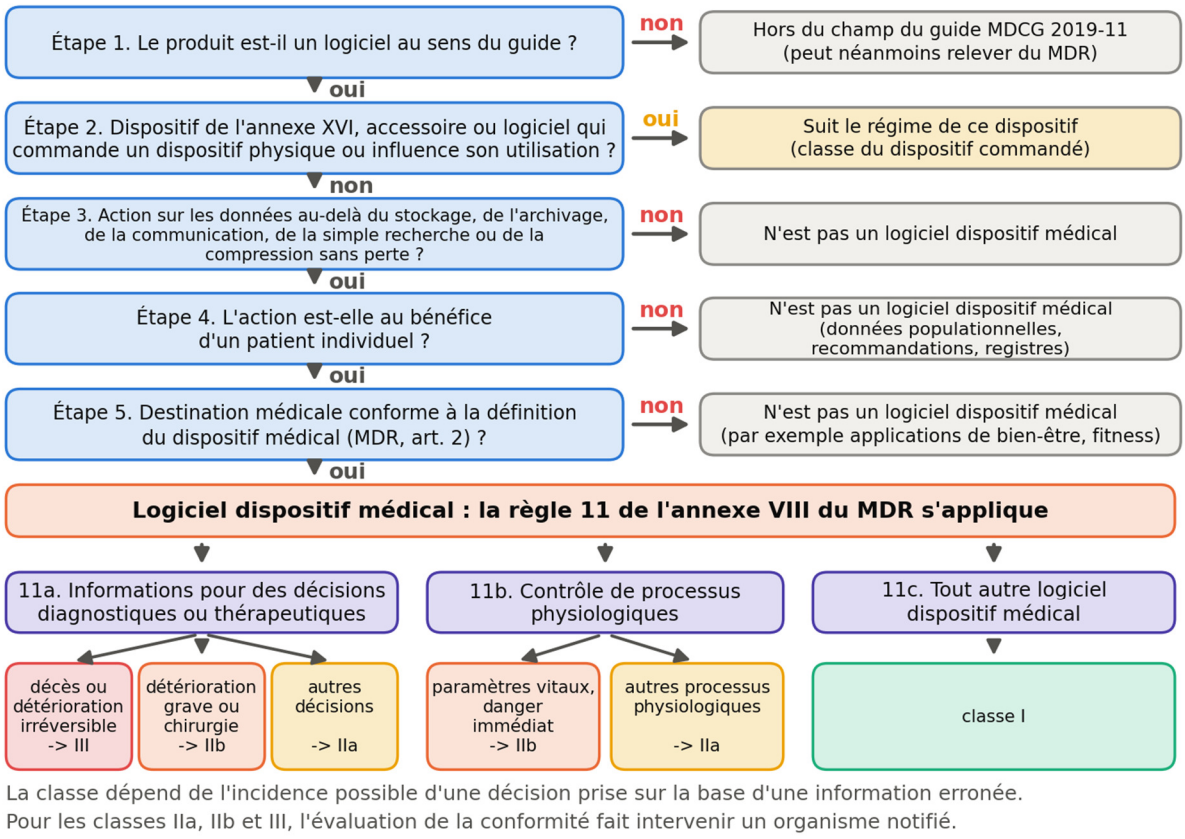


Figure 10.2. Arbre de décision pour la qualification et la classification d'une application comme logiciel dispositif médical, selon les étapes 1 à 5 du guide MDCG 2019-11 rev.1 et la règle 11 de l'annexe VIII du MDR. Le schéma est didactique et ne remplace pas l'évaluation de la conformité. Élaboré d'après [284,372].

10.3.2 La classification selon la règle 11

La règle 11 de l'annexe VIII du MDR comporte trois sous-règles [372]. La sous-règle 11a concerne les logiciels destinés à fournir des informations utilisées pour prendre des décisions à des fins diagnostiques ou thérapeutiques. Ils relèvent de la classe IIa, mais passent en classe IIb si une décision erronée peut entraîner une grave détérioration de l'état de santé ou une intervention chirurgicale, et en classe III si elle peut entraîner la mort ou une détérioration irréversible. La sous-règle 11b concerne les logiciels destinés à contrôler des processus physiologiques, qui relèvent de la classe IIa, sauf lorsqu'ils sont destinés à contrôler des paramètres physiologiques vitaux dont les variations peuvent présenter un danger immédiat pour le patient (classe IIb). Les paramètres vitaux énumérés par le guide sont la respiration, la fréquence cardiaque, les fonctions cérébrales, les gaz du sang, la pression artérielle et la température corporelle. La sous-règle 11c place tous les autres logiciels dispositifs médicaux dans la classe I.

Tableau 10.2. Exemples de classification de logiciels dispositifs médicaux tirés de l'annexe IV du guide MDCG 2019-11 rev.1. Source : [372].

Destination déclarée	Classe	Règle
analyse d'images pour la décision thérapeutique dans l'accident vasculaire cérébral aigu	III	11a
application qui analyse les battements cardiaques, détecte des anomalies et informe le médecin à des fins de diagnostic	IIb	11a
score diagnostique de la dépression à partir des symptômes saisis	IIb	11a
surveillance continue des fonctions vitales en anesthésie ou en soins intensifs	IIb	11b
surveillance à domicile de certains signes vitaux lors de contrôles de routine	IIa	11b
hiérarchisation des options de chimiothérapie pour le médecin	IIa	11a
application de fertilité avec indicateur rouge, vert ou jaune	I	11c
application qui convertit des symboles en parole pour les troubles de la communication	I	11c

La classe détermine la voie vers le marquage CE. Pour les classes IIa et IIb, la procédure d'évaluation de la conformité fait obligatoirement intervenir un organisme notifié, qui vérifie la conception, la production et la surveillance après commercialisation [389]. Une même destination médicale peut conduire à des classes différentes si le contexte change : le guide précise que ses exemples sont indicatifs et que tout changement de la destination ou de la situation clinique peut modifier la classification [372]. Lorsque l'application utilise l'apprentissage automatique, les obligations du règlement sur l'IA (AI Act) s'y ajoutent. Après la modification introduite en 2026 par le règlement (UE) 2026/1744 (Digital Omnibus), les obligations relatives aux systèmes d'IA à haut risque intégrés dans des produits, donc aussi dans les dispositifs médicaux, s'appliquent à partir du 2 août 2028. Les détails figurent au chapitre 13 [22,197].

10.3.3 Les applications remboursées : le modèle allemand des DiGA

L'Allemagne a été le premier pays à intégrer les applications de santé prescrites par le médecin dans la prise en charge de l'assurance maladie obligatoire, à partir de septembre 2020 [390]. Les applications numériques de santé (*Digitale Gesundheitsanwendungen*, DiGA) sont des dispositifs médicaux de classe I ou IIa portant le marquage CE, évalués par l'Institut fédéral des médicaments et des dispositifs médicaux (BfArM) dans le cadre d'une procédure accélérée de trois mois au plus. Elles peuvent être prescrites par des médecins ou des psychothérapeutes et sont remboursées pour environ 73 millions d'assurés [391]. Le fabricant doit démontrer un effet positif sur les soins : un bénéfice médical ou une amélioration de la structure et des processus de soins pertinente pour le patient [392]. Depuis le 26 mars 2024, en vertu de la loi allemande sur la numérisation (DigiG), les dispositifs de classe IIb peuvent eux aussi devenir des DiGA. Pour ceux-ci, l'inscription provisoire en phase d'essai n'est pas prévue, de sorte que le bénéfice médical doit être démontré dès le dépôt de la demande [389].

Les analyses indépendantes montrent à la fois les progrès et les limites du système. Au 1er juillet 2024, 56 DiGA étaient inscrites, 35 à titre permanent et 21 à titre provisoire. Le prix médian initial d'une prescription de trois mois était de 514 euros, ramené après négociation à une médiane de 221 euros, et le nombre de prescriptions par période de déclaration est passé d'environ 41 000 à 209 000 [392]. Les 35 applications inscrites à titre permanent ont toutes démontré leur effet au moyen d'essais randomisés [392]. La qualité de ces essais a toutefois été critiquée. Pour six DiGA des domaines neurologique et psychique, le risque de biais était élevé, principalement en raison de l'absence de mise en aveugle, et dans la plupart des essais l'abandon était plus important dans les groupes d'intervention [390].

10.3.4 Les applications de bien-être dans l'EHDS

L'EHDS introduit une nouvelle obligation pour les applications de bien-être qui déclarent être interopérables avec un système de dossier de santé électronique. Elles doivent porter un label délivré par le fabricant, qui indique les catégories de données pour lesquelles la conformité aux exigences essentielles et aux spécifications communes a été confirmée. Le label est valable trois ans au plus et, pour les applications constituées uniquement de logiciels, il s'affiche sous forme numérique dans l'application (article 47) [23]. En vertu de l'article 105, ces dispositions s'appliquent à partir du 26 mars 2029 pour les premières catégories prioritaires de données. Le label concerne l'interopérabilité et n'équivaut pas au marquage CE d'un dispositif médical.

10.4 L'interopérabilité des dispositifs

10.4.1 La famille ISO/IEEE 11073

Les données d'un dispositif ne sont utiles que si le système qui les reçoit sait ce que signifie chaque valeur. La famille de normes ISO/IEEE 11073 résout ce problème grâce à une nomenclature commune, définie dans la norme de base ISO/IEEE 11073-10101 et étendue par d'autres parties de la famille. Un exemple bien délimité est la partie 10103, republiée par l'ISO en 2025, qui étend la nomenclature aux dispositifs cardiaques implantables : stimulateurs, défibrillateurs, dispositifs de resynchronisation et moniteurs implantables [225]. La norme définit les termes nécessaires pour transmettre un résumé cliniquement pertinent d'une interrogation du dispositif, avec des identifiants qui commencent par le préfixe `MDC_IDC_`. La finalité déclarée est l'échange de données entre les systèmes propriétaires d'interrogation des fabricants et le dossier électronique de la clinique, et la nomenclature peut être utilisée conjointement avec les standards HL7 [225]. En pratique, un rapport de suivi à distance d'un défibrillateur peut arriver dans le dossier du patient sous forme de données discrètes, et non de document numérisé.

10.4.2 Les profils Bluetooth pour la santé

De nombreux dispositifs à usage personnel (tensiomètres, glucomètres, balances, montres) communiquent avec le téléphone par Bluetooth. En juin 2023, l'organisation qui gère le standard, le Bluetooth SIG, a adopté le profil générique GHSP (*Generic Health Sensor Profile*), destiné à différents types de dispositifs de santé personnels, qui transmet les observations de santé à l'aide du modèle de contenu ACOM défini dans la norme IEEE 11073-10206 [383]. L'avantage annoncé est que de nouveaux types de capteurs peuvent être ajoutés sans modifier l'infrastructure qui reçoit les données.

10.4.3 FHIR et les données générées par le patient

Les données générées par le patient (*patient-generated health data*, PGHD) sont des données recueillies par le patient ou par ses dispositifs, en dehors du contact avec le service médical. Pour entrer dans le dossier électronique, elles doivent être traduites en ressources HL7 FHIR, décrites au chapitre 8. Le guide d'implémentation HL7 pour les dispositifs de santé personnels (*Personal Health Device FHIR IG*), version 2.0.0 d'avril 2026, repose sur le modèle ACOM de la norme IEEE 11073-10206 et accepte les données transmises par Bluetooth GHS [384]. Une étude de cas menée avec une montre Garmin a examiné la mise en correspondance du pouls, des pas, du sommeil et du niveau d'activité avec les ressources Observation, Device et Patient, ainsi que la faisabilité de leur transfert vers des serveurs FHIR [385].

La mise en correspondance technique ne règle pas à elle seule le problème de la confiance. Les cliniciens veulent savoir dans quel contexte une mesure a été réalisée et avec quel appareil, autrement dit sa provenance. Un modèle fondé sur des ontologies a proposé de saisir ces informations dès le moment de la collecte et de les transmettre, avec la valeur, au format FHIR [393]. La sécurité constitue la deuxième couche. Une équipe de l'Universitatea Ștefan cel Mare din Suceava (université Ștefan cel Mare de Suceava) a décrit en 2026 une architecture de type *zero trust*, avec des certificats liés au dispositif, un TLS mutuel et une couche d'adaptateurs qui convertit les messages HL7 v2, IEEE 11073 et les formats propriétaires en ressources FHIR. Les auteurs précisent que leurs résultats constituent une validation initiale, et non un test d'intrusion complet [394]. La base juridique du traitement reste celle décrite au chapitre 5, car les données concernant la santé sont une catégorie particulière de données à caractère personnel [118].

10.5 Les grands modèles de langage en consultation

10.5.1 Qu'est-ce qu'un grand modèle de langage ?

Un grand modèle de langage (*large language model*, LLM) est un réseau de neurones entraîné sur de très grands volumes de texte pour prédire le fragment de mot suivant dans une séquence. De cette tâche simple découle la capacité de répondre à des questions, de résumer et de rédiger. Le modèle ne consulte pas une base de faits vérifiés et ne "sait" pas quand il se trompe. L'architecture, l'entraînement et les méthodes d'évaluation sont traités au chapitre 13. Ce qui importe ici, c'est l'utilisation en consultation, par le médecin et par le patient.

10.5.2 La performance : des examens aux cas cliniques

Une méta-analyse de 45 études menées dans plusieurs pays a estimé, aux examens d'autorisation d'exercice de la médecine, une exactitude de 81 % pour GPT-4 et de 58 % pour GPT-3.5. GPT-4 a réussi l'examen dans 26 des 29 situations analysées [395]. Med-PaLM 2 a atteint 86,5 % sur le jeu de questions MedQA, et les médecins évaluateurs ont préféré ses réponses à celles d'autres médecins sur huit des neuf axes cliniques [396]. Une revue de 2025 portant sur 39 jeux de test a replacé ces résultats dans leur contexte. Sur les tests de connaissances, les modèles les plus avancés atteignent 84-90 %, mais sur les tests qui simulent la pratique clinique la performance tombe à 45-69 %, et sur l'évaluation de la sécurité à 40-50 % [397].

Les études qui ont placé le modèle aux côtés du médecin donnent des résultats plus nuancés. Une étude transversale de 2023 a comparé 195 réponses données par des médecins sur un forum public aux réponses de ChatGPT aux mêmes questions. Les évaluateurs ont préféré la réponse du chatbot dans 78,6 % des évaluations, l'ont jugée de bonne ou de très bonne qualité 3,6 fois plus souvent et empathique 9,8 fois plus souvent [398]. Les réponses des médecins comptaient toutefois en moyenne 52 mots, celles du chatbot 211, et le contexte était un forum, non un cabinet.

Dans un essai randomisé portant sur 50 médecins, l'accès à un LLM en plus des ressources habituelles n'a pas amélioré le raisonnement diagnostique : les scores ont été de 76 % et de 74 %, une différence non statistiquement significative. Le modèle utilisé seul a obtenu 16 points de pourcentage de plus que le groupe disposant des ressources conventionnelles [399]. En revanche, dans un essai randomisé portant sur 92 médecins et des cas de prise en charge thérapeutique, le groupe assisté par GPT-4 a obtenu des scores supérieurs de 6,5 points de pourcentage, mais a passé en moyenne près de deux minutes de plus par cas [400]. Un système conversationnel expérimental a eu une meilleure exactitude diagnostique que 20 médecins de famille dans 159 scénarios avec des patients simulés. La consultation s'est toutefois déroulée par chat écrit, un

format inhabituel dans la pratique, et les auteurs appellent à la prudence dans l'interprétation [401]. Sur 2 400 cas réels de pathologie abdominale aiguë, les modèles testés ont moins bien diagnostiqué que les médecins, n'ont pas respecté les recommandations de diagnostic et de traitement et n'ont pas interprété correctement les analyses de laboratoire. Les résultats dépendaient aussi de l'ordre dans lequel les informations leur étaient fournies [402].

10.5.3 Le modèle utilisé par le patient

Le résultat le plus important pour la pratique de 2026 provient d'un essai randomisé préenregistré, publié dans *Nature Medicine* [403]. Au total, 1 298 adultes du Royaume-Uni ont reçu dix scénarios médicaux et ont dû identifier l'affection probable et choisir la conduite à tenir appropriée, de "je reste à la maison" à "j'appelle une ambulance". Ils ont utilisé soit l'un des trois LLM, soit n'importe quelle source qu'ils auraient consultée habituellement. La figure 10.3 résume les résultats.

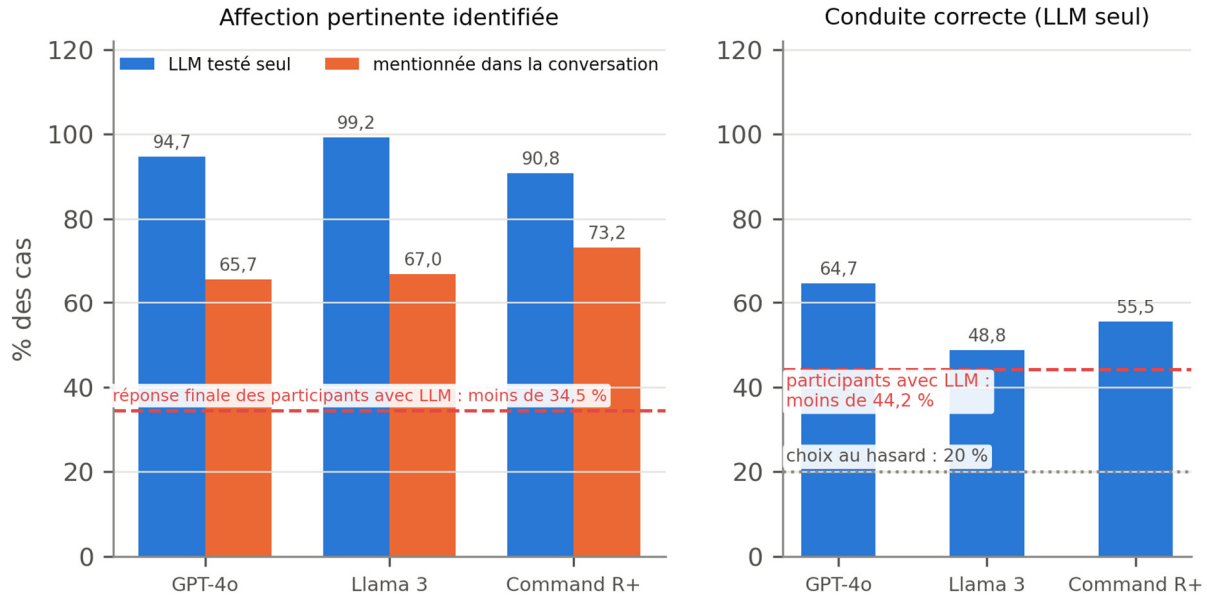


Figure 10.3. Performance de trois modèles de langage utilisés seuls et en conversation avec des utilisateurs ordinaires, dans dix scénarios médicaux. Les barres indiquent les pourcentages rapportés pour chaque modèle. Les lignes horizontales marquent les limites supérieures rapportées pour les réponses finales des participants qui ont utilisé un LLM (moins de 34,5 % et moins de 44,2 %) et le niveau du hasard pour un choix entre cinq conduites (20 %). Données tirées de [403].

Testés seuls, les modèles ont identifié une affection pertinente dans 94,9 % des cas. Les participants qui les ont utilisés l'ont identifiée dans moins de 34,5 % des cas et ont choisi la bonne conduite dans moins de 44,2 %. Leurs résultats n'ont pas été meilleurs que ceux du groupe témoin, qui a même eu des chances 1,76 fois plus élevées de nommer une affection pertinente [403]. L'analyse des conversations a montré d'où venait la différence. Dans 16 des 30 interactions analysées, le premier message de l'utilisateur ne contenait qu'une partie des informations, et les suggestions correctes du modèle n'étaient pas toujours retenues dans la réponse finale [403]. Un test structuré d'un outil de triage destiné aux consommateurs, lancé en janvier 2026, a mis en évidence un sous-triage dans 52 % des urgences de référence : des cas d'acidocétose diabétique ou d'insuffisance respiratoire imminente étaient orientés vers une évaluation dans les 24 à 48 heures [404]. Un audit de cinq chatbots populaires sur des sujets propices à la désinformation a trouvé des réponses problématiques dans 49,6 % des cas. Les chatbots n'ont refusé de répondre que 2 fois sur 250 [405].

10.5.4 Hallucinations, complaisance et références fabriquées

L'hallucination est une affirmation présentée comme un fait, mais qui n'est étayée ni par la source ni par la réalité. Dans un cadre d'évaluation des notes cliniques générées automatiquement à partir des transcriptions des consultations, 12 999 phrases de notes ont été vérifiées par des cliniciens. Le taux d'hallucinations a été de 1,47 % des phrases des notes, et celui des omissions de 3,45 % des phrases des transcriptions [406]. Parmi les phrases comportant des hallucinations, 44 % contenaient des erreurs majeures, qui auraient pu modifier le diagnostic ou la conduite à tenir si elles n'avaient pas été corrigées. Par type, les hallucinations ont été des fabrications (43 %), des négations (30 %), des erreurs de contexte (17 %) et de causalité (10 %), et les fabrications sont apparues surtout dans la section du plan thérapeutique [406].

Une deuxième vulnérabilité est la complaisance (*sycophancy*) : le modèle tend à accepter la prémisse erronée de l'utilisateur. Lorsqu'on leur a demandé des textes partant d'une relation fausse entre deux médicaments équivalents, cinq modèles parmi les plus avancés s'y sont conformés dans des proportions allant jusqu'à 100 % [407]. Le troisième problème concerne la bibliographie. Une étude portant sur 636 références a montré que 55 % de celles générées par GPT-3.5 et 18 % de celles générées par GPT-4 n'existaient pas [408]. La version plus récente s'est trompée moins souvent, mais elle n'a pas pour autant supprimé la nécessité d'une vérification dans PubMed.

10.5.5 Ce qu'un chatbot ne peut pas faire

Le modèle ne travaille qu'avec le texte qu'il reçoit. Il ne voit pas le patient, ne palpe pas l'abdomen, n'observe ni la pâleur ni l'effort respiratoire et ne peut pas demander une analyse. L'étude portant sur des cas d'abdomen aigu a montré que les résultats se dégradent lorsque les informations sont incomplètes ou arrivent dans un autre ordre [402], et l'étude menée avec des utilisateurs ordinaires a montré que ceux-ci omettent souvent des informations nécessaires à une recommandation correcte [403]. Les signes d'alarme (*red flags*) se découvrent par l'anamnèse conduite par le médecin et par l'examen clinique.

10.5.6 La confidentialité des données saisies

Le texte saisi dans un chatbot public se retrouve sur les serveurs du fournisseur. Si ce texte contient des données permettant d'identifier le patient (nom, date de naissance, CNP, adresse, association d'une maladie rare et d'une petite ville), le médecin traite des données concernant la santé, une catégorie particulière protégée par l'article 9 du RGPD [118]. Un tel traitement exige une base juridique et, dans le cas d'un fournisseur externe, des garanties contractuelles. Le médecin doit vérifier si l'outil utilisé les offre, et un compte créé pour un usage privé ne les garantit pas. La confidentialité figure constamment parmi les préoccupations éthiques recensées par une revue systématique de l'utilisation des LLM en médecine [409]. La règle pratique est simple : dans les outils publics, on ne saisit que des cas fictifs ou des données réellement anonymisées, et les outils approuvés par l'établissement s'utilisent conformément à ses procédures.

10.5.7 Lignes directrices, responsabilité et déclaration

L'OMS a publié en janvier 2024 des orientations sur l'éthique et la gouvernance des grands modèles multimodaux, avec plus de 40 recommandations à l'intention des gouvernements, des entreprises technologiques et des prestataires de soins de santé [410]. Ces orientations citent comme risques les affirmations fausses, inexactes ou incomplètes, les données d'entraînement biaisées et le biais d'automatisation (*automation bias*), c'est-à-dire la tendance des professionnels et des patients à ne plus remarquer les erreurs du système [410,411]. La revue systématique

consacrée à l'éthique des LLM a retrouvé la même exigence récurrente : la supervision humaine [409].

La responsabilité de la décision clinique reste celle du médecin, quel que soit l'outil consulté. Dans la rédaction scientifique, les recommandations de l'ICMJE mises à jour en janvier 2026 exigent que les auteurs déclarent l'utilisation d'outils d'IA et décrivent la manière dont ils les ont utilisés. L'aide à la rédaction est mentionnée dans la section des remerciements, tandis que l'utilisation pour la collecte ou l'analyse des données ou pour les figures est décrite dans les méthodes, avec l'outil, la version et, le cas échéant, les instructions données [202]. Les chatbots ne peuvent pas être auteurs, car ils ne peuvent pas répondre de l'exactitude et de l'intégrité du travail. Le contenu généré par l'IA ne peut pas être cité comme source primaire, et les références doivent être vérifiées dans une source bibliographique telle que PubMed [202].

Tableau 10.3. Liste de vérification pour une utilisation responsable d'un LLM par le médecin ou l'étudiant. Source : [118,202,406,408,411].

Étape	Question de contrôle
Avant	Ai-je supprimé toute donnée susceptible d'identifier le patient ?
Avant	L'outil est-il approuvé par l'établissement pour les données cliniques ?
La question	Ai-je fourni le contexte complet (âge, sexe, symptômes, durée, traitement) ?
La réponse	Ai-je vérifié chaque affirmation dans une recommandation ou dans une source primaire ?
La réponse	Ai-je recherché les omissions, en particulier les signes d'alarme ?
Les références	Ai-je trouvé chaque référence dans PubMed et lu le résumé ?
La décision	La décision m'appartient-elle et puis-je la justifier sans le chatbot ?
Après	Ai-je déclaré l'utilisation (outil, version, finalité) ?

10.6 La situation en Roumanie en 2026

Les données d'Eurostat montrent une adoption plus lente des services numériques de santé. En 2025, 42,7 % des personnes en Roumanie ont recherché des informations relatives à la santé sur Internet, contre 60,3 % en moyenne dans l'UE-27. En 2024, 9,1 % ont pris rendez-vous chez un médecin par un site web, contre 39,8 % dans l'UE-27 [297]. L'évolution des prises de rendez-vous en ligne et de l'accès au dossier personnel est présentée au chapitre 8.

Le 1er septembre 2026, la Casa Națională de Asigurări de Sănătate (Caisse nationale d'assurance maladie, CNAS) a lancé le portail eSănătateaMea. L'accès se fait au moyen de l'identité numérique ROeID, et l'assuré peut y consulter ses ordonnances des 5 dernières années, son historique de santé des 10 dernières années et les prestataires auprès desquels il peut prendre rendez-vous en ligne [324]. Le communiqué ne mentionne pas encore de lien avec les dispositifs portables ou avec les applications du patient.

Les applications et les dispositifs portables qui sont des dispositifs médicaux suivent en Roumanie les mêmes règles du MDR que dans le reste de l'Union. Les incidents liés aux dispositifs médicaux sont signalés à l'Agenția Națională a Medicamentului și a Dispozitivelor Medicale din România (Agence nationale du médicament et des dispositifs médicaux de

Roumanie, ANMDMR), dont la page dédiée comporte une rubrique de signalement des incidents [412]. Le règlement EHDS s'applique directement, sans loi de transposition, y compris les dispositions relatives à l'étiquetage des applications de bien-être [23].

Pièges fréquents

- Croire qu'une application est sûre parce qu'elle compte beaucoup de téléchargements ou d'étoiles. La qualité s'évalue avec des instruments validés, tels que MARS.
- Traiter la notification de rythme irrégulier de la montre comme un diagnostic de fibrillation atriale. C'est un signal qui demande une confirmation par ECG.
- Lire la SpO2 affichée par la montre comme une oxymétrie de pouls clinique, sans tenir compte des erreurs et de la pigmentation de la peau.
- Supposer que toute application de santé est un dispositif médical ou, à l'inverse, qu'aucune ne l'est. C'est la destination déclarée par le fabricant qui tranche.
- Confondre le label EHDS d'interopérabilité d'une application de bien-être avec le marquage CE d'un dispositif médical.
- Extrapoler les scores obtenus par les LLM aux examens à leur sécurité en consultation.
- Saisir dans un chatbot public un cas réel "seulement avec les initiales".
- Recopier une référence proposée par un chatbot sans la rechercher dans PubMed.

À retenir

- Les applications mobiles ont des effets réels, mais modestes, les plus constants portant sur l'HbA1c et la pression artérielle.
- L'abandon est la règle : l'observance moyenne des applications destinées aux maladies chroniques a été d'environ 56 %.
- L'exactitude varie beaucoup d'un capteur à l'autre. Le pouls et la détection de la fibrillation atriale sont bien validés, la pression artérielle sans brassard ne l'est pas.
- La destination déclarée par le fabricant détermine si une application est un dispositif médical, et la règle 11 détermine sa classe.
- ISO/IEEE 11073, Bluetooth GHS et HL7 FHIR forment la chaîne par laquelle une mesure parvient au dossier électronique.
- Un LLM excellent aux examens peut n'être que d'une aide limitée pour un patient qui l'utilise seul.
- Vérifiez chaque affirmation et chaque référence, protégez les données du patient et déclarez l'utilisation de l'IA.

Questions d'autoévaluation

1. (QCS) Selon le guide MDCG 2019-11 rev.1, une application de fitness qui compte les pas et les calories, sans aucune destination médicale déclarée :

a) est un dispositif médical de classe I
b) est un dispositif médical de classe IIa, parce qu'elle contrôle un processus physiologique
c) n'est pas qualifiée de logiciel dispositif médical
d) devient un dispositif médical si elle compte plus d'un million d'utilisateurs
e) est réglementée comme un accessoire du téléphone mobile

2. (QCS) Selon la règle 11 du MDR, un logiciel destiné à la surveillance continue des fonctions vitales en soins intensifs relève de la classe :

a) I
b) IIa
c) IIb
d) III
e) ce n'est pas un dispositif médical

3. (QCS) Dans l'essai randomisé de Bean et collaborateurs (2026), les participants issus de la population générale qui ont utilisé un LLM :

a) ont identifié l'affection pertinente dans plus de 90 % des cas
b) ont obtenu de meilleurs résultats que le groupe témoin pour le choix de la conduite à tenir
c) ont identifié l'affection pertinente dans moins de 34,5 % des cas, sans avantage par rapport au groupe témoin
d) ont obtenu les mêmes résultats que les modèles testés seuls
e) ont été exclus de l'analyse finale

4. (QCM) Lesquelles des affirmations sur les dispositifs portables sont étayées par les données du chapitre ?

a) dans une méta-analyse de 2025, les montres connectées ont eu une sensibilité cumulée de 95 % pour la fibrillation atriale
b) seuls 11 % environ des dispositifs portables commercialisés ont été validés pour au moins une mesure
c) la pression artérielle mesurée sans brassard est recommandée pour le diagnostic de l'hypertension
d) les oxymètres de pouls ont dépassé le seuil d'erreur de 3 % dans tous les groupes de pigmentation de la peau
e) la mesure continue du glucose n'a pas modifié l'HbA1c dans les essais randomisés

5. (QCM) Selon les recommandations de l'ICMJE mises à jour en janvier 2026 :

a) l'utilisation de l'IA pour l'aide à la rédaction est déclarée dans la section des remerciements
b) un chatbot peut être coauteur s'il a écrit une partie importante du texte
c) l'utilisation de l'IA pour l'analyse des données est décrite dans la section des méthodes
d) les auteurs répondent du contenu produit à l'aide de l'IA
e) le contenu généré par l'IA peut être cité comme source primaire

CHAPITRE 11

Chapitre 11. Le système d'information, la sécurité de l'information et la signature électronique

Objectifs

À la fin du chapitre, l'étudiant sera capable :

- de distinguer le système d'information d'un établissement de santé du système informatique qui le soutient et d'énumérer leurs composantes
- de décrire les niveaux stratégique, tactique et opérationnel du management d'un système d'information ainsi que les étapes du cycle de vie d'un système informatique
- d'estimer le niveau d'un risque de sécurité à l'aide d'une matrice probabilité x impact, à partir de l'actif, de la menace et de la vulnérabilité
- d'expliquer la différence entre le chiffrement symétrique, le chiffrement asymétrique et la fonction de hachage, et de suivre les étapes de la création et de la vérification d'une signature numérique
- de classer une signature électronique comme simple, avancée ou qualifiée selon le règlement eIDAS, et de reconnaître ses applications dans le système de santé roumain
- d'identifier les organismes de normalisation en informatique médicale et de vérifier sur iso.org le stade actuel d'une norme

11.1 Le système d'information et le système informatique

11.1.1 Deux notions que l'on confond

Dans le langage courant, "le système informatique de l'hôpital" désigne le logiciel installé sur les ordinateurs. L'informatique médicale distingue pourtant deux notions. Un système est un ensemble de personnes, de choses, d'événements et des relations qui les unissent, formant un tout, et un système composé d'éléments humains et techniques est dit sociotechnique [413]. Le système d'information d'un établissement de santé englobe l'ensemble du traitement des données, de l'information et des connaissances au sein de cet établissement, ainsi que les acteurs humains ou techniques qui l'assurent [413]. La partie soutenue par l'ordinateur constitue le système informatique. Le reste demeure non informatisé : le téléphone, le formulaire imprimé, la discussion lors de la relève de garde [413].

Un hôpital possède un système d'information dès son premier jour de fonctionnement, même sans aucun ordinateur. La question est de savoir dans quelle mesure il fonctionne bien et comment on peut l'améliorer [413].

11.1.2 Les composantes

Le système d'information peut être envisagé à travers quatre composantes (figure 11.1).

- **Les personnes** : médecins, infirmières, patients, proches, personnel administratif, informaticiens. Elles interviennent dans le système en tant qu'agents de traitement de l'information, selon des règles établies [413].
- **Les processus** : les fonctions décrivent ce que l'on fait (l'admission, la sortie), sans début ni fin définis, tandis que les processus de travail décrivent comment on le fait, sous la forme d'une succession d'activités avec leurs conditions [413].
- **Les données** : les types d'entités sur lesquelles des données sont conservées, comme "patient", "épisode de soins" ou "lettre médicale" [413].
- **La technologie** : les systèmes d'application (un logiciel installé et configuré pour un établissement donné) et les systèmes physiques de traitement, des serveurs et tablettes jusqu'aux formulaires papier et aux téléphones [413].

La finalité de l'ensemble peut se résumer par la logistique de l'information : la bonne information, au bon moment, au bon endroit, pour les bonnes personnes et sous la bonne forme, afin que celles-ci prennent des décisions correctes [413].

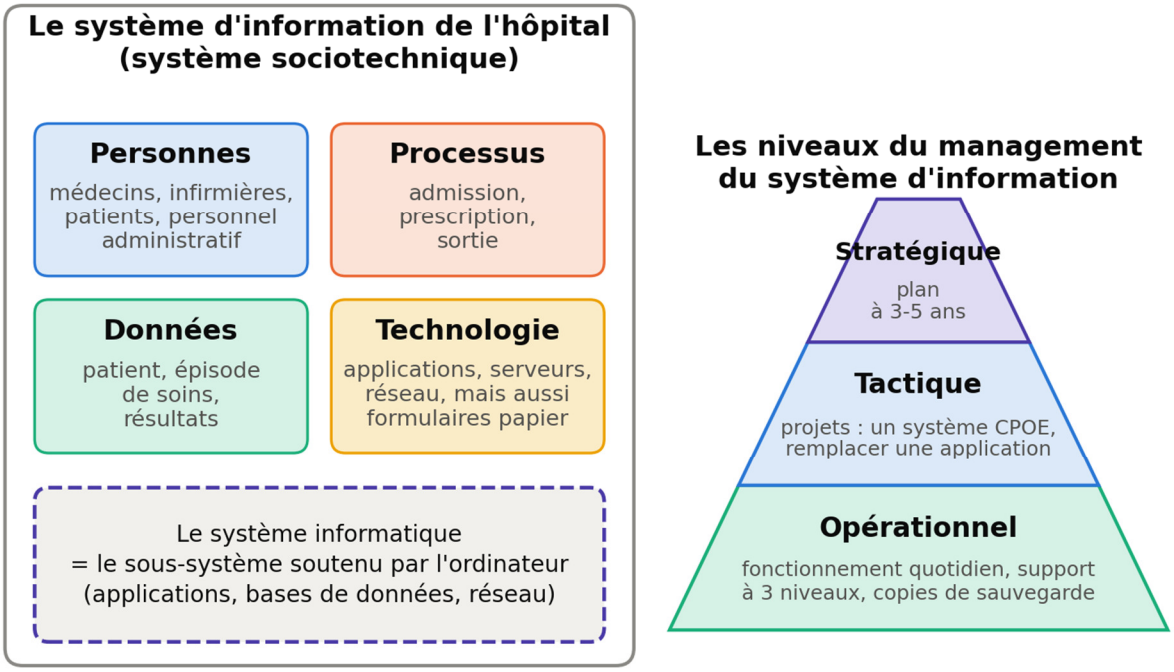


Figure 11.1. À gauche : les composantes du système d'information d'un hôpital, dans lequel le système informatique n'est que le sous-système soutenu par l'ordinateur. À droite : les niveaux du management du système d'information, avec des exemples. Schéma réalisé d'après [413,414].

11.1.3 Les niveaux : stratégique, tactique, opérationnel

Le management des systèmes d'information consiste à planifier le système et son architecture, à diriger sa construction et son fonctionnement sur la base de ces plans et à contrôler leur respect [413]. Il se répartit sur trois niveaux [414].

Le management stratégique considère le système dans son ensemble et produit le plan stratégique de management de l'information. Ce plan décrit les objectifs de l'établissement, la situation actuelle, la situation souhaitée et les étapes permettant de passer de l'une à l'autre ; il est généralement valable de 3 à 5 ans et s'accompagne d'un portefeuille annuel de projets actualisé plus souvent [414]. Le management tactique s'occupe de composantes concrètes que l'on introduit, modifie ou retire, presque toujours au moyen de projets. Des exemples typiques sont l'introduction

d'un système de prescription informatisée (*computerized provider order entry*, CPOE) ou l'évaluation de la satisfaction des utilisateurs dans un service d'anesthésie-réanimation [414]. Le management opérationnel assure le fonctionnement quotidien. Il comprend le service d'assistance aux utilisateurs, le support à trois niveaux (le premier niveau prend en charge l'incident, le deuxième est assuré par les informaticiens de l'hôpital, le troisième, par exemple, par les spécialistes de l'éditeur de l'application) et les copies de sauvegarde [414].

11.1.4 Le cycle de vie d'un système informatique

Un système informatique clinique ne s'"achète" pas une fois pour toutes. D'un point de vue didactique, son cycle de vie peut être divisé en six étapes.

1. **La planification** : le projet part du plan stratégique, avec un plan propre qui précise le problème, les objectifs, les tâches et les ressources [414].
2. **L'analyse des besoins** : on établit ce que le système doit faire (exigences fonctionnelles, par exemple "prescrire et vérifier les interactions") et comment il doit le faire (exigences non fonctionnelles : disponibilité, temps de réponse, sécurité). Pour un système de laboratoire de diagnostic génétique, l'analyse a combiné une revue de la littérature, des entretiens avec les utilisateurs et une analyse du marché, démarche exigée à la fois par le règlement européen relatif aux dispositifs médicaux de diagnostic in vitro et par la norme IEC 62304 [415].
3. **La sélection ou le développement** du logiciel et sa personnalisation pour l'établissement.
4. **La mise en œuvre** : migration des données, formation, mise en service. À la fin, le projet remet le système au management opérationnel au moyen d'un procès-verbal de remise [414].
5. **L'exploitation et la maintenance** : les incidents, les problèmes et les changements techniques sont gérés selon la bibliothèque ITIL, qui recommande un comité chargé de coordonner et de suivre les changements, car toute modification peut avoir des effets secondaires [414].
6. **L'évaluation et le retrait** : l'évaluation indique quand le système doit être remplacé, mais les données des patients doivent parfois être conservées jusqu'à 30 ans, plus longtemps que ne vit n'importe quelle application ; c'est pourquoi l'archivage est confié à un système dédié [131].

11.1.5 Le management du changement

Le terme a deux sens. Au sens technique (ITIL, gestion des changements), le management du changement contrôle les modifications apportées aux composantes du système [414]. Au sens organisationnel (conduite du changement), il concerne les personnes qui doivent travailler autrement après l'introduction du système. Une revue de 90 études sur la mise en œuvre du dossier de santé électronique (*electronic health record*, EHR) a regroupé les facteurs en processus de mise en conformité, de collaboration et de développement des compétences, et en coûts [416]. Une synthèse de 41 études qualitatives ou à méthodes mixtes a montré que la présentation claire de l'objectif et l'implication précoce des utilisateurs préparent la mise en œuvre, tandis qu'une formation adaptée et un support technique à la demande soutiennent l'utilisation après le lancement [417].

La situation en Roumanie en 2026

- La plateforme informatique de l'assurance maladie (Platforma informatică a asigurărilor de sănătate, PIAS) réunit le SIUI, le système de prescription électronique (SIPE), la carte de santé et le dossier électronique de santé (DES) [322].
- Le projet PNRR (plan national de relance et de résilience) de modernisation de la PIAS a été achevé le 31 août 2026. Parmi les résultats, la CNAS cite la migration de la plateforme existante, le portail eSănătateaMea et les formulaires médicaux électroniques [323].
- Depuis le 1er septembre 2026, la composante des formulaires médicaux électroniques (lettres d'orientation [bilete de trimitere], lettres médicales, certificats de congé de maladie) est en service. Pour que les formulaires soient visibles dans le portail eSănătateaMea, les applications des prestataires doivent être adaptées, et la CNAS a recommandé aux prestataires de les mettre à jour conformément aux spécifications d'interfaçage du SIUI [418].

11.2 Le management de la sécurité de l'information

11.2.1 Le système de management de la sécurité de l'information

La protection des données, abordée sous l'angle juridique et éthique au chapitre 5, exige une organisation interne. La norme ISO/IEC 27001:2022 (3e édition, avec l'amendement de 2024) fixe les exigences relatives à un système de management de la sécurité de l'information (SMSI ; *information security management system*, ISMS), qu'un organisme établit, met en œuvre, tient à jour et améliore en continu, avec un processus de management du risque adapté à sa taille [419]. Les mesures concrètes (les contrôles) sont décrites dans l'ISO/IEC 27002:2022.

Pour la santé, l'ISO 27799:2025, 3e édition parue en décembre 2025, adapte ces mesures aux organismes de santé. Son domaine d'application comprend, outre les équipements informatiques habituels, les systèmes EHR, les dispositifs médicaux comportant un logiciel et même les équipements de contrôle des infections ou de gestion technique du bâtiment [189]. Le projet de norme (ISO/DIS 27799:2025) conserve la structure de l'ISO/IEC 27002 en quatre thèmes : mesures organisationnelles, mesures liées aux personnes, mesures physiques et mesures technologiques [190]. La spécificité de la santé tient aux dispositifs médicaux, à l'équilibre entre sécurité clinique et sécurité de l'information, à la très grande diversité des utilisateurs et aux services qui fonctionnent 24 heures sur 24 [190]. La responsabilité de la sécurité incombe au management du système d'information, et la direction de l'établissement doit fournir les ressources financières [131].

11.2.2 L'appréciation du risque

Les mesures ne se choisissent pas dans un catalogue, mais à la suite d'une appréciation du risque dont le domaine est clairement défini [190]. Le guide ISO/IEC 27005:2022 décrit ce processus pour la sécurité de l'information [420]. Le vocabulaire de base comporte cinq termes :

- **l'actif** est ce qui doit être protégé : la base de données des dossiers des patients, le serveur PACS, l'ordinateur portable du médecin de garde, mais aussi la réputation de l'hôpital.
- **la menace** est l'événement susceptible de causer un préjudice (un courriel d'hameçonnage, ou *phishing*, un rançongiciel, une inondation).
- **la vulnérabilité** est la faiblesse par laquelle la menace agit (un mot de passe réutilisé, un système d'exploitation non mis à jour, l'absence de copies de sauvegarde).

- **l'impact** est la conséquence pour les patients et pour l'activité.
- **la probabilité** est la chance que la menace exploite la vulnérabilité dans un intervalle de temps donné.

Le risque combine la probabilité et l'impact. Dans la variante didactique de la figure 11.2, chacun est noté de 1 à 5, et leur produit situe le risque dans une zone de couleur. Les décisions concernant les risques identifiés sont prises selon les critères d'acceptation du risque et les options de traitement définis par l'organisme [190]. En pratique, les risques de la zone critique exigent des mesures immédiates (traitement), tandis que ceux de la zone faible peuvent être acceptés en connaissance de cause. Une revue rédigée pour des non-spécialistes a constaté que les stratégies de protection employées par les hôpitaux des États-Unis sont souvent faibles [421].

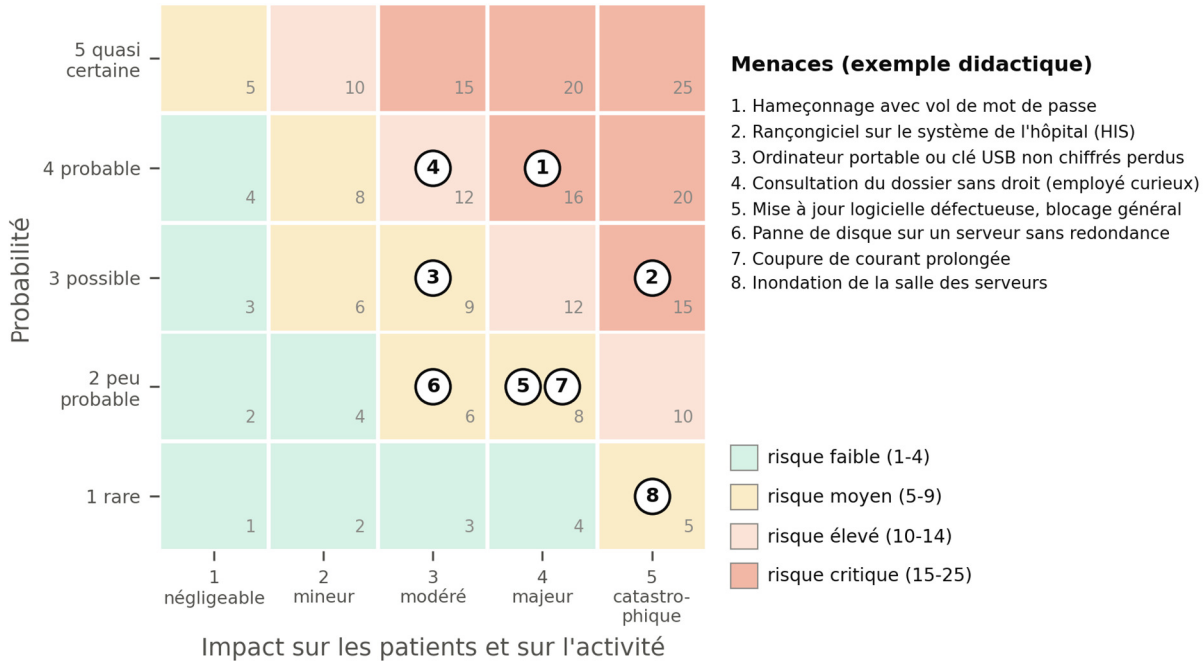


Figure 11.2. Matrice de risque probabilité x impact pour huit menaces typiques d'un hôpital. Les positions sont illustratives, choisies à des fins didactiques, et ne proviennent pas d'une appréciation réelle. Chaque hôpital établit ses propres scores.

11.2.3 Politiques, classification de l'information et contrôle d'accès

La politique de sécurité est le document approuvé par la direction qui indique ce qui est permis et qui en répond. Elle s'appuie sur la classification de l'information. Le projet ISO 27799 recommande de classer les informations personnelles de santé au minimum comme "confidentielles" et admet des degrés différents : une fracture de jambe n'a pas la même sensibilité qu'une infection sexuellement transmissible, un épisode psychiatrique survenu il y a dix ans diffère d'un épisode actuel, et les données génétiques concernent aussi les apparentés du patient [190].

Le contrôle d'accès fondé sur les rôles (*role-based access control*, RBAC) attribue des droits au rôle (médecin du service, infirmier, pharmacien, agent d'accueil), et non à la personne. Un utilisateur ne devrait accéder aux données d'un patient que lorsqu'il fait partie de l'équipe qui le soigne ou qu'il agit en son nom ; pour les urgences, il faut une procédure autorisée de contournement des règles habituelles, appelée "bris de glace" (*break the glass*) [190]. Un contrôle trop strict peut empêcher, en urgence, la prescription d'un médicament par un médecin d'un autre

service. C'est pourquoi, après une analyse du risque, certaines restrictions peuvent être assouplies pour le personnel médical, en contrepartie d'une responsabilité individuelle accrue [414].

11.2.4 Les journaux d'audit

Chaque consultation du dossier électronique laisse une trace dans le journal d'audit (*audit log*) : qui, quoi, quand, depuis quel poste. Une revue systématique de 85 études a montré que ces journaux, conçus au départ pour le contrôle d'accès, servent aujourd'hui aussi à la recherche sur l'activité clinique [422]. Dans les litiges pour faute médicale, le journal d'audit peut montrer qui a vu un résultat et à quel moment [423], raison pour laquelle le journal lui-même doit être protégé [190]. Le composant européen de journalisation exigé des systèmes EHR par le règlement EHDS est présenté au chapitre 5.

11.2.5 La continuité d'activité et la reprise

L'hôpital doit fonctionner même quand le système informatique ne fonctionne pas. Une revue a recensé 43 pannes dans 166 hôpitaux des États-Unis en 2012-2018, totalisant 701 jours, dont près de la moitié (48,8 %) étaient liées à des cyberattaques [424]. Une interruption n'est pas toujours due à une attaque. Le 19 juillet 2024, une mise à jour défectueuse d'un logiciel de sécurité a été suivie d'une perte de réponse aux analyses (*scans*) effectuées depuis Internet dans 759 des 2 232 hôpitaux américains analysés (34,0 %) [425].

Le plan de continuité d'activité identifie les processus vitaux et les procédures de secours (formulaires papier, tableurs). Le projet ISO 27799 exige que les plans soient testés par paliers, des exercices sur table jusqu'aux répétitions complètes [190]. Le plan de reprise repose sur des copies de sauvegarde elles-mêmes protégées contre les attaques et sur des centres de données comportant au moins deux sites redondants [131]. La notification des incidents selon la directive NIS2 et l'OUG n° 155/2024, de même que l'attaque par rançongiciel de 2024 contre des hôpitaux roumains, sont analysées au chapitre 5.

11.3 Notions de cryptographie pour le médecin

11.3.1 Le chiffrement symétrique

Le chiffrement transforme un texte en clair en un texte chiffré, qui ne peut être relu qu'avec la clé appropriée. Dans le chiffrement symétrique, l'expéditeur et le destinataire utilisent la même clé secrète. La norme actuelle est l'AES (*Advanced Encryption Standard*), qui traite les données par blocs de 128 bits, avec des clés de 128, 192 ou 256 bits [426]. Le chiffrement symétrique demande peu de calcul et convient aux grands volumes : le disque d'un ordinateur portable, l'archive PACS, les copies de sauvegarde. Son problème est la distribution de la clé. Comment la clé parvient-elle au destinataire sans être interceptée ?

11.3.2 Le chiffrement asymétrique

La cryptographie à clé publique utilise une paire de clés liées mathématiquement. La clé publique peut être connue de tous, la clé privée reste uniquement chez son titulaire. Ce qui est chiffré avec la clé publique du destinataire ne peut être déchiffré qu'avec sa clé privée. Inversement, seul le détenteur de la clé privée peut générer une signature numérique, et n'importe qui peut la vérifier avec la clé publique [427]. La norme américaine de signature numérique (FIPS 186-5) approuve trois familles d'algorithmes (RSA, ECDSA et EdDSA) et exige pour RSA un

module d'au moins 2 048 bits [427]. En pratique, le chiffrement est hybride : le document est chiffré de façon symétrique, et seule la clé symétrique, courte, est chiffrée de façon asymétrique.

11.3.3 La fonction de hachage

Une fonction de hachage (*hash*) calcule, à partir de données de longueur quelconque, une valeur de longueur fixe [428]. La même entrée donne toujours la même empreinte, et la modification d'une seule virgule produit une empreinte complètement différente. La fonction de hachage n'est pas un chiffrement : elle n'a pas de clé, et le document ne peut pas être reconstitué à partir de l'empreinte. Son rôle est de vérifier l'intégrité : lors de la signature, c'est en règle générale l'empreinte, version condensée du document, qui est signée, et non le document entier [427].

11.3.4 Les certificats numériques et l'infrastructure de clé publique

Comment le pharmacien sait-il qu'une clé publique appartient au médecin X ? Grâce au certificat numérique, un document électronique par lequel une autorité de certification (AC ; *certification authority*, CA) lie la clé publique à l'identité de son titulaire. Les autorités sont organisées de manière hiérarchique, de l'AC racine aux AC intermédiaires, et l'ensemble forme l'infrastructure de clé publique (ICP ; *public key infrastructure*, PKI). L'ISO 17090-1:2021 décrit pour la santé les certificats d'identité, les certificats d'attributs (par exemple, le rôle professionnel) et les hiérarchies d'AC [429]. L'ISO 17090-4:2026, publiée en janvier 2026, fixe les exigences minimales et les formats pour la génération et la vérification des signatures numériques apposées sur les documents de santé [430]. La version de projet de cette partie décrit les formats européens CAdES, XAdES, PAdES et JAdES et leurs niveaux, parmi lesquels la signature de base, la signature horodatée et la signature d'archivage, qui permet la vérification après de nombreuses années. La clé privée du médecin est conservée sur une carte à puce, sur un jeton USB ou sur un jeton logiciel [428].

11.3.5 TLS, HTTPS et le chiffrement de bout en bout

Lorsque l'adresse d'un portail commence par https, la connexion est protégée par le protocole TLS (*Transport Layer Security*). TLS permet aux applications de communiquer par Internet d'une manière conçue pour empêcher l'interception, la modification et la falsification des messages. La version 1.3 a été republiée en juillet 2026 sous la référence RFC 9846, qui remplace la RFC 8446 [431]. Le serveur présente son certificat, le navigateur vérifie la chaîne de certification, puis une clé de session symétrique est établie. TLS ne protège les données que pendant leur transport.

Dans le chiffrement de bout en bout (*end-to-end*), le message n'est déchiffré que sur l'appareil du destinataire, et le serveur intermédiaire ne peut pas le lire. Aucune technique n'est toutefois utile si les données partent par la messagerie personnelle, une habitude également documentée au chapitre 5.

On ne s'attend pas à ce que les algorithmes RSA et ECDSA résistent à l'attaque d'un ordinateur quantique de grande taille [427]. En août 2024, le NIST a publié la norme post-quantique FIPS 203, qui décrit le mécanisme ML-KEM d'établissement de clés, considéré actuellement comme sûr même face à un attaquant disposant d'un ordinateur quantique [432].

Exemple

Exemple didactique, avec de petits nombres (le RSA réel utilise des nombres de plus de 600 chiffres décimaux). Le médecin possède la clé publique (n = 55, e = 3) et la clé privée d = 27.

- L'empreinte fictive de la lettre médicale est h = 9.
- La signature : s = h^d mod n = 9^27 mod 55 = 4. Le médecin envoie la lettre et la signature s = 4.
- La vérification chez le destinataire : s^e mod n = 4^3 mod 55 = 64 mod 55 = 9. Le destinataire recalcule l'empreinte de la lettre reçue, obtient 9, et les valeurs coïncident : la lettre n'a pas été modifiée et a été signée avec la clé privée correspondante.
- Si quelqu'un modifie la dose dans la lettre, l'empreinte recalculée devient, par exemple, 10. La vérification donne toujours 9, et la signature est rejetée.

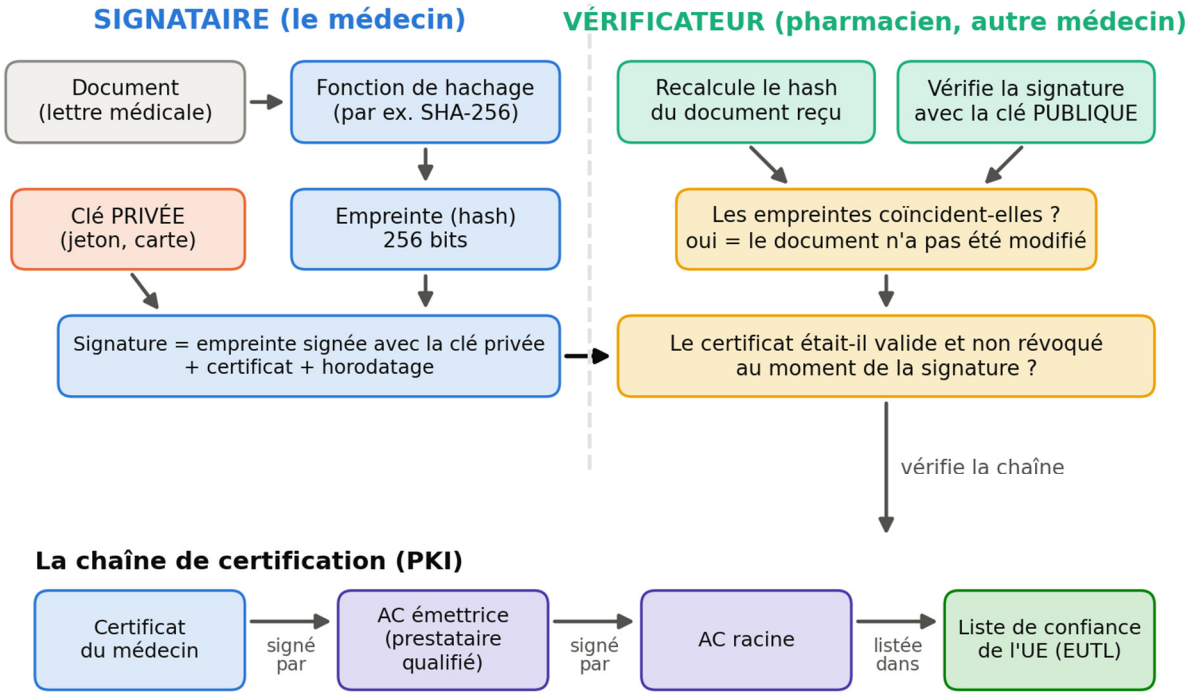


Figure 11.3. Création et vérification d'une signature numérique : signature de l'empreinte (hash) avec la clé privée, vérification avec la clé publique du certificat et remontée de la chaîne de certification jusqu'à la liste de confiance de l'UE. Schéma réalisé d'après [427,428,433].

11.4 La signature électronique

11.4.1 Les types de signature selon eIDAS

Le règlement (UE) n° 910/2014 (eIDAS) définit trois niveaux (tableau 11.1). La signature électronique (simple) désigne des données sous forme électronique, qui sont jointes ou associées logiquement à d'autres données sous forme électronique et que le signataire utilise pour signer [433]. La signature électronique avancée doit satisfaire à quatre exigences : être liée au signataire de manière univoque, permettre d'identifier le signataire, avoir été créée à l'aide de données que le signataire peut, avec un niveau de confiance élevé, utiliser sous son contrôle exclusif, et être liée aux données associées à cette signature de telle sorte que toute modification ultérieure des données soit détectable [433]. La signature électronique qualifiée est une signature électronique avancée créée à l'aide d'un dispositif de création de signature électronique qualifié et reposant sur un certificat qualifié [433].

Les effets juridiques diffèrent. L'effet juridique d'une signature électronique ne peut être refusé au seul motif qu'elle se présente sous une forme électronique ou qu'elle n'est pas qualifiée, mais seule la signature qualifiée a un effet juridique équivalent à celui d'une signature manuscrite [433]. Depuis la modification de 2024, la reconnaissance de la signature qualifiée dans tous les États membres est prévue par le nouvel art. 24 bis [434]. Les personnes morales (un hôpital, une caisse d'assurance maladie) utilisent le cachet électronique, qui garantit l'origine et l'intégrité des données [433].

Tableau 11.1. Les types de signature électronique selon le règlement eIDAS. Source : [433,434].

Type	Exigences principales	Effet juridique	Exemples
Simple	données électroniques utilisées par le signataire pour signer	l'effet ne peut être refusé au seul motif de la forme électronique	nom tapé au clavier, image scannée de la signature
Avancée	liée au signataire de manière univoque, l'identifie, sous son contrôle exclusif, rend les modifications détectables	comme pour la signature simple, l'effet ne peut être refusé au seul motif de la forme électronique	signature avec un certificat non qualifié
Qualifiée	avancée + certificat qualifié + dispositif qualifié de création	équivalente à la signature manuscrite, reconnue dans tous les États de l'UE	signature du médecin sur jeton ou carte d'un prestataire qualifié

11.4.2 L'horodatage et la validation

L'horodatage électronique lie un ensemble de données à un instant précis et prouve que ces données existaient à ce moment. L'horodatage électronique qualifié bénéficie d'une présomption d'exactitude de la date et de l'heure ainsi que d'intégrité des données, et sa source de temps est liée au temps universel coordonné (UTC) [433]. Il est important parce qu'un certificat peut expirer ou être révoqué, et la signature ne reste valable que s'il est prouvé qu'elle a été créée avant.

La validation d'une signature qualifiée confirme, entre autres, que le certificat était qualifié et valide au moment de la signature, qu'il avait été délivré par un prestataire de services de confiance qualifié et que l'intégrité des données signées n'a pas été compromise [433]. Chaque État membre publie une liste de confiance des prestataires qualifiés qu'il contrôle [433], et la Commission européenne propose un outil en ligne de consultation de ces listes (TP 11). À long terme, les signatures électroniques qualifiées ne peuvent garantir l'intégrité des documents archivés que si le système d'archivage est capable de renouveler les signatures et les algorithmes de hachage devenus obsolètes [131].

11.4.3 Applications dans le système de santé roumain

Depuis le 1er février 2014, les médicaments du traitement ambulatoire, avec ou sans contribution personnelle, sont prescrits sur le formulaire de prescription médicale électronique, en ligne ou hors ligne, qui comporte une partie remplie par le médecin et une autre par le pharmacien (arrêté MS/CNAS n° 674/252/2012). Lorsque le médecin dispose d'une signature électronique étendue (semnătură electronică extinsă, terme de la législation antérieure à 2024), le document porte la mention qu'il a été généré et signé électroniquement. Le médecin ne signe de sa main que s'il ne dispose pas d'une signature électronique étendue, s'il ne peut pas l'utiliser ou en mode hors ligne [435]. Dans sa version consolidée au 29 juillet 2026, cette mention renvoie encore à la loi n° 455/2001 [435], alors que cette loi a été abrogée par la loi n° 214/2024 relative à l'utilisation de la signature électronique, de l'horodatage et à la prestation de services de confiance, en vigueur depuis octobre 2024. La nouvelle loi désigne l'Autorité pour la numérisation de la Roumanie

(Autoritatea pentru Digitalizarea României, ADR) comme organe de contrôle des services de confiance [436].

Depuis le 1er septembre 2026, la composante CNAS des formulaires médicaux électroniques est en service ; elle comprend notamment les lettres d'orientation vers les spécialités ou en vue d'une hospitalisation, les bons d'examens paracliniques, la lettre médicale, les prescriptions de soins à domicile et le certificat de congé de maladie. Le patient peut les consulter dans le portail eSănătateaMea si l'application du prestataire a été adaptée [418]. Dans le portail, le patient s'authentifie par ROeID [324]. La lettre médicale de sortie, une fois archivée, reste vérifiable à long terme si le système d'archivage utilise des signatures électroniques qualifiées et les renouvelle en temps utile [131].

11.4.4 eIDAS 2 et le portefeuille européen d'identité numérique

Le règlement (UE) 2024/1183 a modifié eIDAS et est entré en vigueur le 20 mai 2024. Chaque État membre doit fournir au moins un portefeuille européen d'identité numérique (*European Digital Identity Wallet*), une application dans laquelle le citoyen conserve ses données d'identification personnelle et des attestations électroniques d'attributs ; le code source des composants de l'application doit faire l'objet d'une licence à code source ouvert (*open source*) [434]. La Commission européenne précise que les portefeuilles seront disponibles d'ici la fin de 2026 [437].

11.5 La normalisation en informatique médicale

11.5.1 Pourquoi des normes et des standards

Les standards d'interopérabilité réduisent le coût des interfaces entre applications, tandis que la conversion (le mappage) d'un format à l'autre fait généralement perdre de l'information ou du sens [131]. HL7 version 2 est le standard de communication le plus implémenté dans les hôpitaux, DICOM couvre l'imagerie, et IHE (*Integrating the Healthcare Enterprise*) n'est pas un organisme de normalisation. IHE choisit et restreint les standards existants pour des flux cliniques concrets, au moyen de profils testés lors des Connectathons [131]. La version actuelle de HL7 FHIR est R5 (5.0.0, mars 2023), et R6 n'était pas publiée au 27 septembre 2026 [307]. DICOM existe aussi sous la forme d'une norme ISO, l'ISO 12052:2026 [438]. Le chapitre 8 traite ces standards en détail.

11.5.2 ISO/TC 215 et CEN/TC 251

Le comité technique ISO/TC 215 *Informatique de santé* (*Health informatics*) a été créé en 1998 et son secrétariat est assuré par l'ANSI (États-Unis). Au 27 septembre 2026, sa page indiquait 251 normes ISO publiées (212 sous la responsabilité directe du comité, chiffres qui incluent les mises à jour) et 63 en cours d'élaboration, avec 39 membres participants et 32 membres observateurs [439]. Il compte deux sous-comités (génomique, pharmacie et médicaments) et des groupes de travail consacrés à l'architecture (WG 1), à l'interopérabilité des systèmes et des dispositifs (WG 2), au contenu sémantique (WG 3), à la sécurité, à la sûreté et à la confidentialité (WG 4) ou encore à la télésanté (WG 12), ainsi qu'un groupe commun avec l'ISO/IEC JTC 1/SC 42 pour l'intelligence artificielle [439].

Les données ouvertes de l'ISO permettent de suivre l'évolution dans le temps (figure 11.4). De 2001 au 2 septembre 2026, l'ISO/TC 215 et ses sous-comités ont publié 392 documents (hors amendements). Au 27 septembre 2026, 246 étaient en vigueur (118 normes internationales, 82

spécifications techniques, 45 rapports techniques et une spécification publiquement disponible), et 146 avaient été retirés ou remplacés [438]. Chaque année de la période 2021-2025, le comité a publié de 22 à 25 documents.

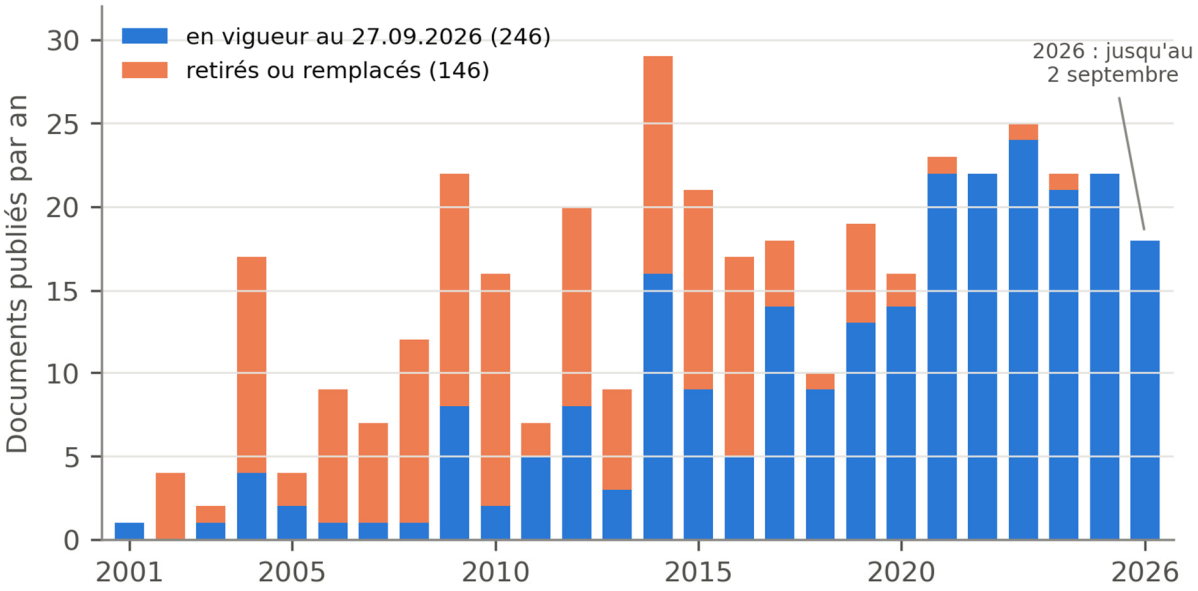


Figure 11.4. Documents publiés par l'ISO/TC 215 et ses sous-comités, par année de publication, selon leur état au 27 septembre 2026. L'année 2026 est incomplète. Calcul de l'auteur sur le jeu de données iso_deliverables_metadata, téléchargé le 27 septembre 2026 (licence ODC-By 1.0), hors amendements et corrigenda [438].

Son homologue européen est le CEN/TC 251 *Informatique de santé* (*Health informatics*). En vertu de l'Accord de Vienne, dans la mesure du possible, une seule norme est approuvée pour chaque sujet, adoptée par l'ISO et par le CEN, et le CEN/TC 251 fait partie du conseil de l'initiative conjointe pour la normalisation mondiale en informatique de santé (dix membres en 2022) [440]. La norme relative au résumé international du patient a démarré comme une initiative menée par le CEN/TC 251, avec un financement de la Commission européenne, et figure parmi les normes du CEN/TC 251 adoptées par l'ISO/TC 215 [440]. L'édition actuelle est l'ISO 27269:2025 [438].

11.5.3 L'organisme national

En Roumanie, les normes européennes et internationales sont adoptées par l'intermédiaire de l'Association roumaine de normalisation (Asociația de Standardizare din România, ASRO). Les normes d'informatique médicale relèvent du comité technique CT 319 "Informatică medicală" (informatique médicale), qui a publié, par exemple, la SR EN ISO/IEEE 11073-10103:2025 relative à la nomenclature des dispositifs cardiaques implantables [441]. L'auteur de cet ouvrage est expert au sein de ce comité technique.

11.5.4 Les normes changent : exemples

Le tableau 11.2 réunit quelques normes liées aux thèmes du chapitre. La colonne "état" montre pourquoi une norme doit toujours être vérifiée sur iso.org avant d'être citée. La série ISO 21549 (les données de la carte de santé du patient) avait fait l'objet de nouvelles éditions en 2023 et 2024, retirées en avril 2025. À l'inverse, les projets ISO/DIS 17090-4, ISO/DIS 22532 et ISO/DIS 16791 sont devenus des normes publiées en 2026 [438].

Tableau 11.2. Normes d'informatique médicale liées à la sécurité, à la signature et à l'identification. État vérifié sur iso.org le 27 septembre 2026. Source : [189,419,430,438].

Norme	Objet	État au 27 septembre 2026
ISO/IEC 27001:2022	exigences relatives aux systèmes de management de la sécurité de l'information	en vigueur (amendement 2024)
ISO 27799:2025	mesures de sécurité de l'information dans le domaine de la santé	en vigueur, 3e édition
ISO 17090-4:2026	signatures numériques pour les documents de soins de santé	en vigueur, 3e édition
ISO 12052:2026	DICOM, flux de travail et gestion des données	en vigueur, 3e édition
ISO 17523:2025	exigences relatives aux prescriptions électroniques	en vigueur, 2e édition
ISO 27269:2025	résumé international du patient	en vigueur, 2e édition
ISO 11239:2023	identification des médicaments : formes pharmaceutiques, voies d'administration, conditionnements	en vigueur, 2e édition
ISO 22532:2026	identification des médicaments (IDMP) : vocabulaire de base	en vigueur, 1re édition
ISO 16791:2026	codage lisible par machine des identifiants de conditionnement	en vigueur, 3e édition
IEC/TS 81001-2-2:2025	communication des besoins, des risques et des mesures de sécurité pour les logiciels de santé	en vigueur, 1re édition
ISO/TR 11636:2009	réseau privé virtuel à la demande pour l'infrastructure d'information de santé	en vigueur
ISO 21549-5:2023 et 21549-7:2024	données d'identification et de médication de la carte de santé	retirées en 2025

11.6 Le facteur humain

11.6.1 L'hameçonnage

L'hameçonnage (*phishing*) est un message qui imite une source de confiance pour obtenir un mot de passe, un clic ou l'ouverture d'une pièce jointe. Une revue systématique de 70 articles sur le facteur humain dans la sécurité des organismes de santé a constaté une croissance constante de la littérature sur l'hameçonnage et a souligné la nécessité d'une hygiène numérique du personnel médical sur les réseaux sociaux [442].

Les études qui mesurent le comportement réel sont plus instructives. Dans un hôpital italien de plus de 6 000 salariés, trois campagnes simulées ont montré que le message personnalisé, lié au contexte local, est beaucoup plus efficace : lors de la première campagne, 55 % des destinataires ont cliqué sur le lien personnalisé, contre 7 % pour le message standard (figure 11.5). Les auteurs décrivent aussi les obstacles : la réaction des syndicats, la difficulté de la coordination et les problèmes éthiques posés par les faux messages [443]. Une autre étude, discutée au chapitre 5, a associé le clic à la charge de travail, et non à l'intention déclarée [181]. Dans un hôpital norvégien, le taux de clics a été de 53 % dans le groupe neutre (396 sur 753), et une brève intervention fondée sur la dissonance cognitive n'a pas produit de différence statistiquement significative (P = 0,22) [444].

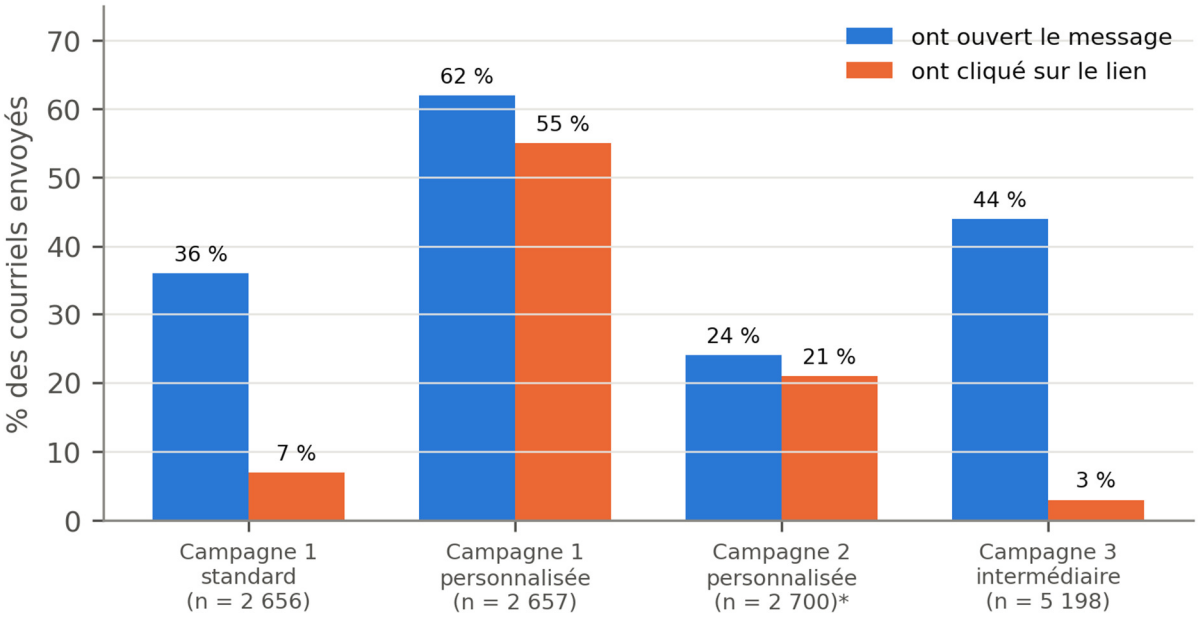


Figure 11.5. Résultats de trois campagnes d'hameçonnage simulé dans un hôpital italien de plus de 6 000 salariés : pourcentage de messages ouverts et de liens consultés, rapporté aux courriels envoyés. *Lors de la campagne 2, 1 600 messages ont été placés dans le dossier des courriers indésirables. Données issues de [443].

11.6.2 Mots de passe et authentification multifacteur

Le guide NIST SP 800-63B-4 (2025) a modifié plusieurs règles apprises par des générations d'utilisateurs. Un mot de passe utilisé comme facteur unique doit compter au moins 15 caractères, et celui qui est utilisé dans une authentification multifacteur au moins 8. On n'impose plus de combinaisons obligatoires de types de caractères ni de changement périodique du mot de passe, mais le mot de passe doit être changé lorsqu'il existe des preuves de compromission, et il est vérifié par rapport à une liste de valeurs compromises ou très utilisées [445]. Les applications évaluées au niveau d'assurance AAL2 doivent aussi proposer une option résistante à l'hameçonnage, et le code envoyé par SMS reste un authentificateur restreint [445].

L'authentification multifacteur (*multi-factor authentication*, MFA) combine quelque chose que l'on sait (le mot de passe), quelque chose que l'on possède (le téléphone, le jeton) et quelque chose que l'on est (l'empreinte digitale). Les patients accèdent à eSănătateaMea par le compte ROeID, avec l'adresse électronique, le mot de passe et le code généré par l'application d'authentification [324].

11.6.3 L'hygiène numérique à l'hôpital

Le guide de l'ENISA de 2025 pour le secteur de la santé regroupe les mesures en huit domaines. On y trouve notamment la MFA pour les systèmes accessibles en ligne, le principe du moindre privilège, au moins une copie de sauvegarde résistante aux rançongiciels ou hors ligne, la politique du bureau propre et de l'écran vide (*clean desk/clear screen*) et les simulations d'hameçonnage [446].

Pour l'étudiant, quelques règles s'appliquent dès le premier jour de stage. On ferme la session ouverte avant de quitter le poste. Le mot de passe ne se prête pas au collègue de garde, car le journal d'audit attribuera les actions au titulaire. Les résultats ne se photographient pas avec le téléphone personnel. Un courriel suspect se signale, il ne se supprime pas en silence. Les incidents de cybersécurité peuvent être signalés à la DNSC par la plateforme PNRISC ou au numéro 1911 [447].

Pièges fréquents

- Croire que le management du changement se limite à l'installation d'une nouvelle version. Sans formation ni implication des utilisateurs, un bon système peut échouer.
- Considérer le hachage comme une forme de chiffrement. La fonction de hachage n'a pas de clé et ne peut pas être inversée ; elle ne cache donc rien, elle vérifie seulement l'intégrité.
- Croire qu'une signature scannée collée dans un PDF est une signature qualifiée. C'est une signature électronique simple.
- Envoyer sa clé privée ou le code PIN de son jeton à un collègue "pour qu'il signe les ordonnances". La signature est attribuée au titulaire.
- Lire le cadenas du navigateur comme une garantie de la légitimité du site. HTTPS protège la connexion, il ne vérifie pas l'honnêteté de celui qui a obtenu le certificat.
- Citer une norme d'après un projet (DIS) ou d'après une édition retirée sans vérification sur iso.org.

À retenir

- Le système informatique est le sous-système informatisé du système d'information, qui comprend aussi les personnes, les processus et les outils papier.
- Le management stratégique produit le plan sur 3 à 5 ans, le management tactique conduit les projets, le management opérationnel assure le fonctionnement quotidien.
- Le risque s'estime comme probabilité x impact, pour chaque triplet actif-menace-vulnérabilité.
- Le chiffrement symétrique utilise une seule clé, le chiffrement asymétrique une paire de clés, et la fonction de hachage produit une empreinte de longueur fixe.
- On signe l'empreinte avec la clé privée, et la vérification se fait avec la clé publique du certificat.
- Seule la signature électronique qualifiée a un effet juridique équivalent à celui de la signature manuscrite.
- Au 27 septembre 2026, l'ISO/TC 215 comptait 246 documents en vigueur, et le stade d'une norme se vérifie sur iso.org.

Questions d'autoévaluation

1. (QCS) Lequel des éléments suivants N'EST PAS une exigence applicable à la signature électronique avancée, selon l'art. 26 du règlement eIDAS ?

a) être liée au signataire de manière univoque
b) permettre d'identifier le signataire
c) avoir été créée à l'aide de données placées sous le contrôle exclusif du signataire
d) permettre de détecter toute modification ultérieure des données signées
e) avoir été créée à l'aide d'un dispositif de création de signature électronique qualifié

2. (QCS) Le certificat de signature d'un médecin a été révoqué le 10 mars. Quel élément permet de prouver qu'une lettre médicale qu'il a signée le 15 janvier est valable ?

a) la fonction de hachage du document
b) l'horodatage appliqué au moment de la signature
c) la clé publique du destinataire
d) le chiffrement du document par AES
e) le certificat du serveur HTTPS

3. (QCS) Dans l'exemple didactique RSA avec la clé publique (n = 55, e = 3), le destinataire reçoit la signature s = 4 et recalcule l'empreinte du document reçu : h = 10. Quelle conclusion est correcte ?

a) la signature est valide, car 4^3 mod 55 = 9 est inférieur à 10
b) la signature ne correspond pas au document reçu, car 4^3 mod 55 = 9 diffère de 10
c) la vérification nécessite la clé privée du médecin
d) l'empreinte doit d'abord être déchiffrée avec la clé symétrique
e) le résultat ne peut pas être interprété sans l'horodatage

4. (QCM) À propos des niveaux du management du système d'information d'un hôpital, les affirmations correctes sont :

a) le management stratégique produit un plan généralement valable de 3 à 5 ans
b) l'introduction d'un système CPOE est un projet typique du management tactique
c) le service d'assistance et le support à trois niveaux relèvent du management opérationnel
d) le management tactique s'occupe exclusivement des équipements matériels
e) le plan stratégique, une fois approuvé, n'est plus actualisé

5. (QCM) Selon le guide NIST SP 800-63B-4, les affirmations correctes sont :

a) un mot de passe utilisé comme facteur unique compte au moins 15 caractères
b) les utilisateurs doivent être obligés de changer leur mot de passe tous les 90 jours
c) on n'impose pas de règles de composition du type "une majuscule, un chiffre, un symbole"
d) le mot de passe nouvellement choisi est vérifié par rapport à une liste de valeurs compromises ou très utilisées
e) le code reçu par SMS est l'authentificateur le plus résistant à l'hameçonnage

CHAPITRE 12

Chapitre 12. Les systèmes d'information hospitaliers et le système informatique unique intégré de l'assurance maladie

Objectifs

À la fin du chapitre, l'étudiant sera capable :

- d'énumérer les modules d'un système d'information hospitalier et de préciser quelles données chacun produit et quelles données il consomme ;
- de comparer l'architecture monolithique et l'architecture modulaire intégrée par un moteur d'interfaces, en calculant le nombre d'interfaces nécessaires dans chaque cas ;
- d'interpréter les données publiées après 2020 sur la prescription informatisée, les systèmes d'aide à la décision clinique et la fatigue face aux alertes ;
- de décrire le parcours des données d'un patient du médecin de famille à la caisse d'assurance maladie, avec le rôle du SIUI, du SIPE, du DES et de la carte de santé ;
- de calculer la durée moyenne de séjour, le taux d'occupation des lits, l'indice de case-mix et le taux de réhospitalisation, puis de les interpréter sur un tableau de bord ;
- d'expliquer ce que signifie le fonctionnement en mode dégradé d'un hôpital lorsque le système informatique devient indisponible.

12.1 Le système d'information de l'hôpital : composantes et modules

12.1.1 Système d'information et système informatique

Tout hôpital possède un système d'information dès son premier jour de fonctionnement. Winter et ses collaborateurs définissent le système d'information d'un établissement de santé comme le sous-système sociotechnique qui englobe l'ensemble du traitement des données, des informations et des connaissances, ainsi que les acteurs humains et techniques qui l'assurent [448]. La feuille de température posée sur la table de chevet, le registre de garde et le téléphone entre le service et le laboratoire en font partie. La partie qui utilise des ordinateurs, que Winter appelle le sous-système d'information informatisé, correspond à ce que le langage courant nomme le système d'information hospitalier (SIH ; *hospital information system*, HIS). La bonne question n'est donc pas de savoir si l'hôpital "a un système", mais comment on peut en améliorer le fonctionnement, par exemple par une gestion systématique ou par des outils modernes [448].

12.1.2 Les modules

Un SIH se compose d'applications spécialisées (composants d'application), chacune soutenant une fonction de l'hôpital. Le tableau 12.1 résume les plus fréquentes.

Le système de gestion administrative des patients, appelé couramment module ADT (*admission, discharge, transfer*), enregistre l'identité, l'admission, les transferts entre services et la

sortie. Il fournit des données administratives correctes et à jour à toutes les autres composantes, raison pour laquelle Winter le décrit comme le centre de la mémoire administrative de l'hôpital [448]. Une erreur d'identification commise à l'admission se propage ensuite au laboratoire, à l'imagerie et à la facturation.

La feuille d'observation clinique générale (foaia de observație clinică generală, FOCG) électronique réunit l'anamnèse, l'examen clinique, l'évolution, l'épicrise et la lettre médicale. Le dossier de soins infirmiers se trouve dans le même espace. La prescription informatisée des médicaments et des examens (*computerized provider order entry*, CPOE) transmet les demandes à la pharmacie, au laboratoire ou à la radiologie et peut proposer la vérification de la dose, des interactions et des allergies [448]. Le système informatique de laboratoire (*laboratory information system*, LIS) reçoit les demandes, assure le suivi des échantillons, récupère les résultats des analyseurs et les valide. Le système d'information de radiologie (RIS) planifie les examens et gère les comptes rendus, tandis que l'archive d'images (PACS) stocke les images DICOM (chapitre 7).

Le bloc opératoire dispose de son propre module de planification des salles et de protocole opératoire. La réanimation utilise un système de gestion des données du patient (*patient data management system*, PDMS), qui récupère automatiquement les paramètres des appareils situés au lit du patient et calcule des scores de gravité tels que le TISS ou le SAPS II [448]. Les modules administratifs (finances et comptabilité, ressources humaines, approvisionnement, pharmacie avec gestion des stocks) forment le progiciel de gestion intégré (*enterprise resource planning*, ERP). L'entrepôt de données (*data warehouse*) extrait périodiquement des données de l'ensemble de ces modules pour les rapports et les tableaux de bord.

Tableau 12.1. Les principaux modules d'un système d'information hospitalier. Source : adapté de [448].

Module	Fonctions principales	Données typiques	Échange de données habituel
ADT (gestion administrative des patients)	identification, admission, transfert, sortie	CNP, données démographiques, service, lit, diagnostic d'adressage	HL7 v2 ADT, FHIR Patient/Encounter
FOCG électronique et soins infirmiers	anamnèse, évolution, épicrise, plan de soins	notes cliniques, diagnostics codés CIM	documents cliniques structurés (chapitre 8)
CPOE avec aide à la décision	prescriptions, demandes d'examens, protocoles de prescription	médicament, dose, voie, fréquence	messages de demande et de résultat
Pharmacie	validation, dispensation, gestion des stocks	consommation par patient et par service	intégration avec le CPOE et l'ERP
LIS	demandes, prélèvement, résultats, validation	valeurs numériques, unités, intervalles de référence	HL7 v2, codes LOINC
RIS/PACS	planification, images, comptes rendus	études DICOM, comptes rendus textuels	DICOM, HL7 (chapitre 7)
Bloc opératoire et réanimation (PDMS)	planification des salles, protocole d'anesthésie, surveillance, scores	signes vitaux à des intervalles de quelques secondes ou minutes, perfusions, bilan hydrique	passerelle pour dispositifs médicaux
Finances, codage DRG, ERP	facturation à l'assurance, coûts, salaires, achats	codes de diagnostic et d'acte, consommations	déclarations aux plateformes nationales

12.1.3 Architectures : monolithique ou modulaire

La manière dont les modules sont reliés entre eux définit l'architecture. Si toutes les applications écrivent dans une seule base de données centrale, l'architecture est de type DB1. Si

chaque application a sa propre base, elle est de type DBn, et les mêmes données (le patient, le séjour hospitalier) existent en plusieurs copies qu'il faut maintenir cohérentes [448]. En pratique, presque tous les grands hôpitaux ont une architecture mixte. Un deuxième critère est le nombre d'applications : l'architecture monolithique utilise un seul composant pour toutes les fonctions, alors que l'architecture modulaire en utilise plusieurs. Le troisième critère est le nombre de fournisseurs. La solution "tout chez un seul fournisseur" (*all-in-one*) s'oppose à la combinaison "le meilleur produit pour chaque fonction" (*best-of-breed*), qui comporte en règle générale plusieurs bases de données [448].

Dans une architecture modulaire, les applications doivent échanger des messages. Si chaque paire d'applications dispose d'une liaison directe, le nombre d'interfaces pour n applications est n(n-1)/2 : 10 applications donnent 45 interfaces, chacune devant être maintenue et testée à chaque mise à jour. La littérature qualifie cette situation d'architecture "spaghetti". La solution classique est un serveur de communication, appelé aussi moteur d'interfaces (*integration engine*), placé au centre d'une architecture en étoile (figure 12.1). Chaque application ne communique qu'avec le serveur ; seules n liaisons sont donc nécessaires [448]. Le serveur reçoit les messages de façon asynchrone, les conserve dans une file d'attente si le destinataire est arrêté, les convertit d'un format à un autre et peut les diffuser simultanément à plusieurs destinataires. Un exemple typique : le module ADT envoie un message HL7 d'admission, et le serveur le retransmet à toutes les applications qui auront besoin des données du patient (par exemple le CPOE et la radiologie), chacune mettant alors à jour sa copie locale des données du patient [448]. Les standards de messagerie (HL7 v2, FHIR) et le dossier de santé électronique sont traités au chapitre 8.

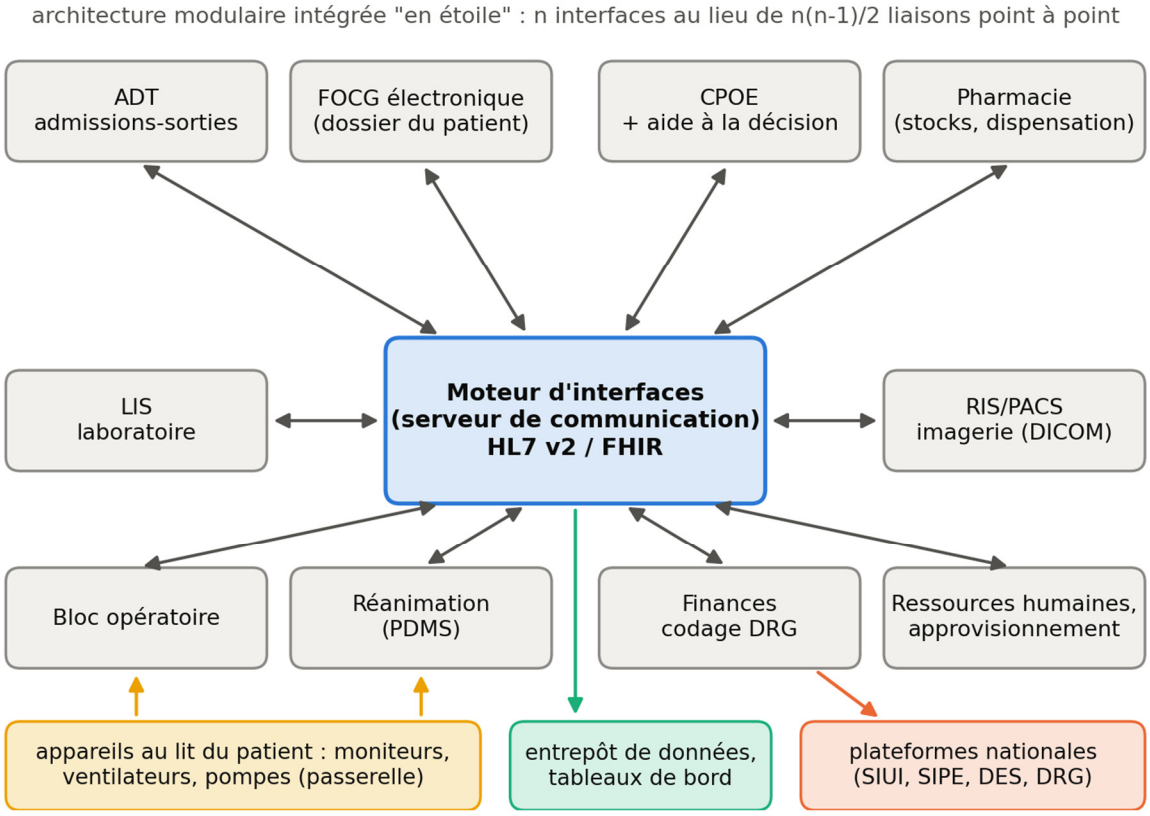


Figure 12.1. Architecture modulaire d'un système d'information hospitalier intégré par un moteur d'interfaces (serveur de communication). Les appareils situés au lit du patient entrent par une passerelle dans les modules du bloc opératoire et de la réanimation. Les données agrégées partent vers l'entrepôt de données et vers les plateformes nationales. Schéma didactique, adapté de [448].

Le choix de l'architecture n'est pas une question purement technique. Dans une analyse portant sur plus de 1 600 hôpitaux allemands du projet DigitalRadar, un recours plus important aux modules d'un même fournisseur et un taux d'intégration plus élevé ont été associés à des scores de maturité numérique plus élevés, sur une échelle de 0 à 100 [449]. Les auteurs recommandent des achats stratégiques de SIH, des améliorations ciblées des systèmes et des incitations de la part des politiques publiques.

12.2 La prescription informatisée à l'hôpital et l'aide à la décision clinique

12.2.1 Ce qu'apporte le CPOE

Sur papier, une prescription passe par plusieurs mains : le médecin écrit, l'infirmière retranscrit sur la feuille de traitement, la pharmacie lit et dispense. Chaque retranscription est une occasion d'erreur. Le CPOE supprime la retranscription, impose des champs obligatoires (dose, voie, fréquence) et permet des vérifications automatiques. Les systèmes d'aide à la décision clinique (*clinical decision support*, CDS) ajoutent le calcul de la dose, la vérification des interactions, des allergies et de la compatibilité du médicament avec les résultats de laboratoire, ainsi que des protocoles de prescription prédéfinis pour les situations fréquentes [448].

Les données de la période 2020-2026 confirment le bénéfice, avec des nuances. Une revue systématique de 35 études hospitalières a comparé la prescription assistée par les technologies de l'information aux prescriptions sur papier. La technologie a été associée à des taux plus faibles d'erreurs de prescription, avec un rapport de cotes médian de 0,24 (intervalle interquartile 0,03-0,57) [450]. La même revue a toutefois trouvé des erreurs générées par la technologie elle-même dans 31 des 35 études, liées à la saisie des données, aux prescriptions en double et à la fatigue face aux alertes. Pour la chimiothérapie, où une erreur de dose peut être fatale, la méta-analyse de huit études a obtenu un risque relatif de 0,19 (IC à 95 % 0,08-0,44) après l'introduction du CPOE, avec une hétérogénéité très élevée (I^2 = 99 %) et une étude dans laquelle les erreurs ont augmenté [451]. Une méta-analyse de 45 études sur le CDS appliqué à la prescription a rapporté un effet global positif, mais faible (différence moyenne standardisée de 0,084, IC à 95 % 0,067-0,102), et, dans certains sous-groupes, la différence par rapport à la pratique habituelle n'était pas statistiquement significative [452]. La figure 12.2 met ces valeurs côte à côte.

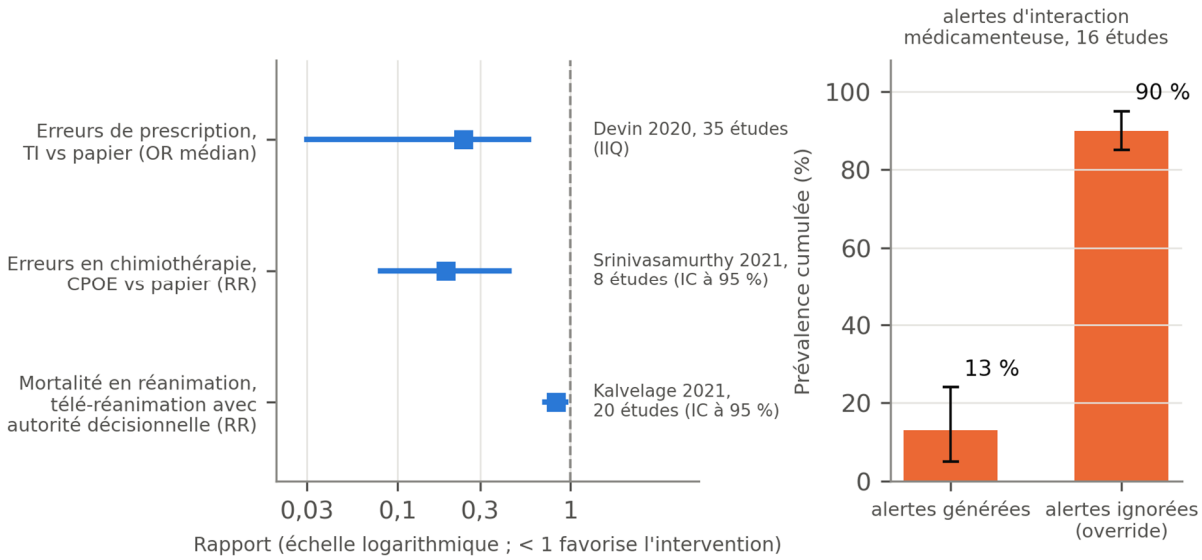


Figure 12.2. Estimations publiées concernant les technologies de l'information appliquées à la prescription et à la réanimation. À gauche : erreurs de prescription à l'hôpital (OR médian, intervalle interquartile, 35 études) [450] ; erreurs médicamenteuses en chimiothérapie après CPOE (RR, IC à 95 %, 8 études) [451] ; mortalité en réanimation avec une télé-réanimation dotée d'une autorité décisionnelle (RR, IC à 95 %, 20 études) [453]. À droite : prévalence cumulée des alertes d'interaction médicamenteuse générées et de celles ignorées par les médecins (IC à 95 %, 16 études) [454]. Valeurs reprises exactement des résumés cités.

12.2.2 La fatigue face aux alertes

Une alerte utile doit apparaître rarement, au bon moment, et demander une action claire. La réalité est différente. Une méta-analyse de 16 études sur les alertes d'interaction médicamenteuse a estimé une prévalence globale des alertes générées de 13 % (IC à 95 % 5-24 %), et les médecins les ignorent (*override*) dans 90 % des cas (IC à 95 % 85-95 %) [454]. Un taux élevé d'alertes ignorées ne signifie pas automatiquement de la négligence. Beaucoup d'alertes ne sont pas pertinentes sur le plan clinique, et le médecin qui en voit des dizaines par jour apprend à les fermer sans les lire. Ce phénomène s'appelle la fatigue face aux alertes (*alert fatigue*).

Sa mesure est étonnamment imprécise. Une revue de 22 revues systématiques n'a trouvé qu'une seule définition opérationnelle de la fatigue face aux alertes, et les indicateurs les plus utilisés étaient le nombre d'alertes, le taux d'alertes ignorées et le taux d'acceptation [331]. Les auteurs proposent de mesurer la fatigue par la baisse durable, par rapport à un niveau de base, du taux de réponse appropriée aux alertes. Du côté des remèdes, les hôpitaux qui ont publié la manière dont ils optimisent leurs alertes ont tous eu recours à un comité pluridisciplinaire, et cinq évaluations sur six ont rapporté une diminution du nombre d'alertes après les modifications [455]. La leçon pour le futur médecin : une alerte se paramètre, se mesure et se retire si elle ne remplit pas son rôle.

12.3 Étude de cas : l'informatisation de la réanimation

12.3.1 Pourquoi la réanimation

Un patient sous ventilation mécanique produit des milliers de valeurs par jour : fréquence cardiaque, pression artérielle invasive, saturation, paramètres du ventilateur, débit de chaque pompe à perfusion. Sur papier, l'infirmière en recopie une partie une fois par heure. Un PDMS les récupère automatiquement sur les appareils, affiche des avertissements, calcule des scores et génère lors du transfert une synthèse pour le service qui prend en charge le patient [448]. Cette abondance de signaux a aussi un coût. Les alarmes fausses ou non pertinentes sur le plan clinique

entraînent une désensibilisation du personnel, et une revue systématique de 69 publications a trouvé des éléments montrant que les solutions informatiques peuvent réduire la fatigue face aux alarmes, la majorité des travaux visant l'élimination des alarmes techniquement fausses [456]. Dans une méta-analyse portant sur 6 908 infirmières de réanimation, la fatigue face aux alarmes a atteint un niveau modéré, et les valeurs plus élevées allaient de pair avec une tendance accrue aux erreurs médicales et avec l'épuisement professionnel [457].

12.3.2 Un cahier des charges roumain

Le ministère de la Santé a publié le cahier des charges du projet "Sistem informatic pentru evidența clinică a secțiilor ATI" (système informatique d'enregistrement clinique des services de réanimation, SIEC-ATI), lancé en mai 2021 et financé par le Programme opérationnel Compétitivité 2014-2020 (code SMIS 131065) [458]. Ce document est utile sur le plan didactique, car il montre comment on formule les exigences d'un système clinique avant qu'un fournisseur soit choisi.

Étude de cas

- Périmètre : 18 hôpitaux et instituts de Bucarest, Cluj-Napoca, Iași, Târgu Mureș et Timișoara, avec 649 lits de réanimation et 275 postes d'anesthésie dans les salles d'opération, plus une composante centrale au ministère [244].
- Intégration des appareils : enregistrement automatique, toutes les 90 secondes au plus, de tous les paramètres du moniteur, du ventilateur, des pompes volumétriques, des pousse-seringues et de l'appareil d'hémofiltration ; la collecte ne doit pas dépendre du modèle d'appareil, afin que le remplacement d'un équipement n'affecte pas le système [244].
- Sécurité médicamenteuse : alertes d'allergie avec un niveau de gravité paramétrable, prescriptions exigeant une double signature, doses inhabituelles signalées séparément et éventuellement bloquées [244].
- Scores et codage : calcul automatique des scores de Glasgow, APACHE II, SAPS II et TISS ; codes CIM-10 et export vers le système de l'hôpital des données nécessaires au DRG, avec la durée de ventilation et le score APACHE [244].
- Interopérabilité : interfaces XML/JSON, HL7/FHIR et services web avec le SIH de chaque hôpital [244].
- Niveau national : un tableau de bord au ministère indiquant les lits de réanimation libres et occupés de chaque hôpital, actualisé en 60 secondes au plus, avec la possibilité de réserver un lit [244].
- Continuité : possibilité de créer temporairement l'enregistrement du patient si le SIH de l'hôpital est indisponible [244].

Il vaut la peine d'observer comment les exigences sont rédigées. Elles sont vérifiables ("90 secondes au maximum", "60 secondes au maximum"), neutres vis-à-vis du fabricant (toute marque citée porte la mention "ou équivalent") et séparent les fonctions cliniques des fonctions d'infrastructure [244]. Un cahier des charges vague ("système moderne, facile à utiliser") ne peut pas être vérifié lors de la réception.

12.3.3 Ce que montrent les études

Une étude rétrospective de type avant-après, réalisée par une équipe de Timișoara dans un service de réanimation traumatologique, a comparé 43 patients de la période de documentation sur papier (2021-2022) à 65 patients de 2025, après l'introduction d'un PDMS. La durée médiane de séjour en réanimation est passée de 13 à 6 jours ($p = 0,02$), mais, dans les analyses ajustées, les

différences de mortalité, d'infections et de durée de séjour n'étaient pas statistiquement significatives [459]. Ce résultat illustre une règle générale : une petite étude avant-après suggère un effet causal, mais ne le démontre pas. À Timişoara également, un résumé consacré à la numérisation des services de réanimation de l'Hôpital clinique départemental d'urgence "Pius Brânzeu" a discuté les bénéfices de la surveillance avancée pour la sécurité du patient, ainsi que les risques liés à la fiabilité des systèmes, à la cybersécurité et à la formation du personnel [460]. La préparation des personnes n'est pas uniforme. Dans une enquête menée auprès de 161 professionnels d'un service de réanimation roumain, les infirmières ont déclaré plus souvent que les médecins ne pas pouvoir apprécier l'utilité des outils numériques, et les simulateurs de formation avaient été utilisés par 1,8 % des infirmières, contre 47,2 % des médecins [461].

La littérature internationale ajoute trois autres axes. Le premier est la télémédecine : une méta-analyse de 20 études portant sur 477 637 patients a trouvé une mortalité en réanimation plus faible lorsque l'équipe de télé-réanimation disposait d'une autorité décisionnelle (RR 0,82, IC à 95 % 0,71-0,94), mais pas lorsqu'elle se limitait à un rôle de conseil (figure 12.2) [453]. Le deuxième est la vérification automatique du respect des recommandations : dans cinq hôpitaux universitaires dotés de systèmes informatiques différents, un système qui appliquait des recommandations codées numériquement à des données standardisées a obtenu une exactitude de 97,0 %, contre 86,6 % pour les évaluateurs humains [462]. Le troisième est la qualité des données. Une revue de 29 études en réanimation a rapporté que les erreurs médicamenteuses liées au dossier électronique représentaient 34 % des erreurs médicamenteuses en réanimation, et que le texte copié-collé apparaissait dans jusqu'à 82 % des notes d'évolution des internes [463]. Une analyse d'ensemble de la numérisation de la réanimation discute aussi le potentiel des modèles d'apprentissage automatique pour surpasser les scores classiques dans la prédiction de la mortalité et de la durée de séjour [464].

12.4 Le système informatique de l'assurance maladie en Roumanie

12.4.1 Les composantes de la plateforme

L'hôpital ne travaille pas de manière isolée. Les services qu'il fournit aux assurés sont déclarés, validés et remboursés par l'intermédiaire d'une plateforme nationale. La loi n° 95/2006, telle que modifiée par la loi n° 45/2019, prévoit que la Caisse nationale d'assurance maladie (Casa Națională de Asigurări de Sănătate, CNAS) organise et administre la plateforme informatique de l'assurance maladie (Platforma informatică din asigurările de sănătate, PIAS). Celle-ci comprend le système informatique unique intégré (SIUI), le système national de la carte d'assurance sociale de santé, le système national de prescription électronique (SIPE) et le système du dossier électronique de santé (DES) [321,322].

Chaque composante a un rôle distinct. Dans le SIUI, les prestataires de services médicaux et pharmaceutiques déclarent par voie électronique les services fournis. Le SIPE a introduit la prescription électronique des ordonnances remboursées et gratuites. Pour la carte électronique d'assurance maladie (CEAS), la CNAS publie des bibliothèques logicielles compatibles également avec la carte d'identité électronique (CEI) [322]. Le DES réunit, au niveau national, des enregistrements électroniques comportant des données cliniques, biologiques, diagnostiques et thérapeutiques accumulées tout au long de la vie. Le dossier est constitué lors de la transmission du premier document médical du patient et comporte cinq modules : "Résumé d'urgence", "Historique médical", "Antécédents déclarés par le patient", "Documents médicaux" et "Données personnelles" [321]. Les médecins ne voient les données qu'avec le consentement du patient.

L'exception est le résumé d'urgence, accessible sans consentement aux médecins des structures d'urgence et des soins primaires, strictement pour l'acte médical ; le patient peut consulter l'historique complet des accès [321]. La structure du dossier et ses standards sont discutés au chapitre 8. Le tableau 12.2 résume le rôle de chaque composante et sa situation au 27 septembre 2026.

Tableau 12.2. Les composantes du système informatique de l'assurance maladie : rôle et situation au 27 septembre 2026. Source : [321-325,465].

Système	Utilisateurs	Rôle	Situation au 27 septembre 2026
SIUI	tous les prestataires sous contrat avec les caisses d'assurance maladie	déclaration des services fournis, base du remboursement	opérationnel, migré sur l'infrastructure modernisée de la PIAS
SIPE	médecins prescripteurs, pharmacies	ordonnances électroniques remboursées et gratuites	opérationnel
Carte d'assurance (CEAS) et CEI	patient, prestataires	identification de l'assuré lors de la fourniture du service	la CEI ne devient pas obligatoire au 1.09.2026, la carte nationale reste utilisée
DES	médecins, patient	historique médical consolidé, résumé d'urgence	à part, le portail eSănătateaMea (1.09.2026) montre à l'assuré son historique de santé sur 10 ans
Déclaration DRG (INMSS)	hôpitaux	ensemble minimal de données par patient, groupage RO DRG	groupeur RO DRG v1 en ligne, application DRGNational 1.40 (31.03.2026)

12.4.2 Le parcours des données d'un patient

Suivons un patient fictif (figure 12.3). Le médecin de famille l'examine, constate une affection cardiaque et émet une lettre d'orientation (bilet de trimitere). En consultation externe, le cardiologue examine le patient et réalise une échographie, puis déclare les services dans le SIUI. Le patient est hospitalisé. À la sortie, l'hôpital émet la lettre médicale et l'ordonnance, et l'ensemble minimal de données au niveau du patient parvient à l'Institut national de management des services de santé (Institutul Național de Management al Serviciilor de Sănătate, INMSS). Celui-ci collecte les données de tous les hôpitaux, pour l'hospitalisation complète et l'hospitalisation de jour, et met à disposition le groupeur RO DRG [465]. Les diagnostics sont codés en ICD-10-AM, 3e édition, et la classification en vigueur est la RO DRG v.1.1, de mars 2024 [139]. Les détails du groupage figurent au chapitre 4. L'ordonnance électronique passe par le SIPE, la pharmacie la dispense, et la caisse départementale d'assurance maladie rembourse les services sur la base des déclarations.

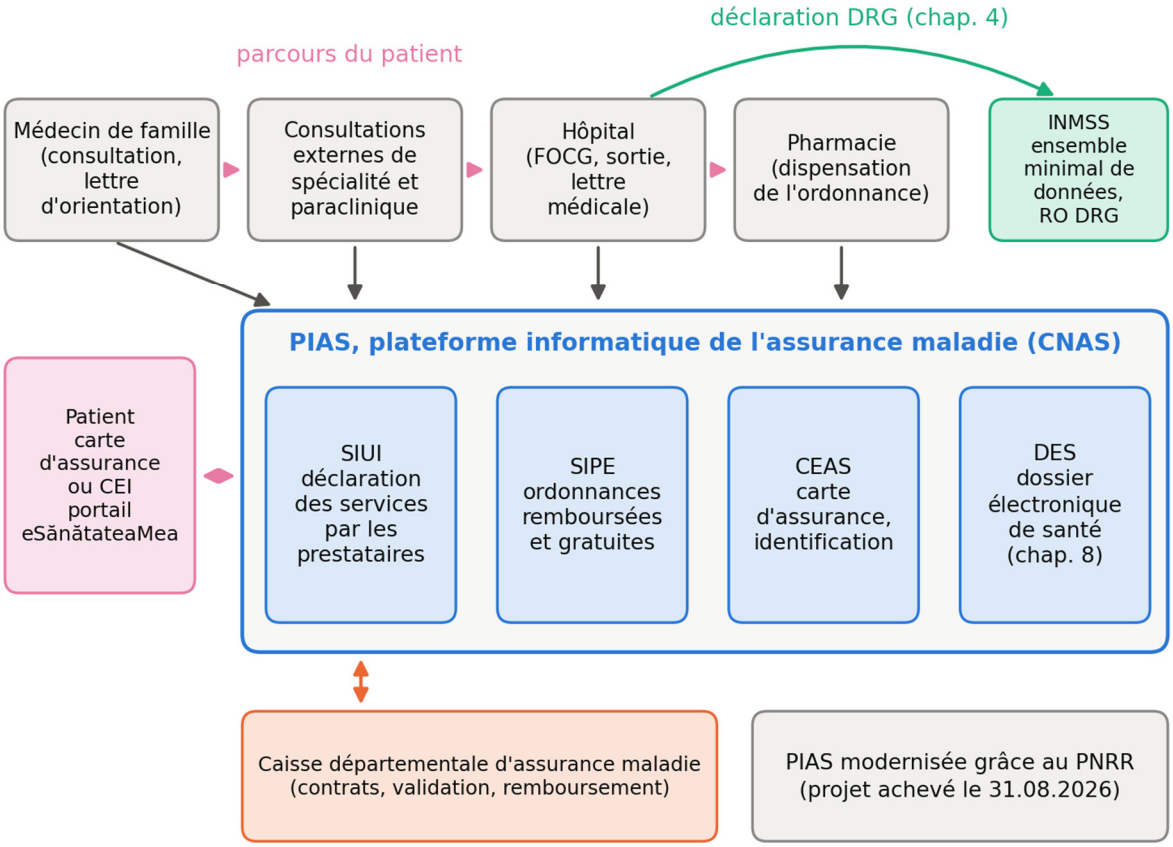


Figure 12.3. Flux de données entre les prestataires et la plateforme informatique de l'assurance maladie (PIAS) : parcours du patient, déclaration au SIUI, prescription par le SIPE, identification par la carte d'assurance ou la CEI, dossier électronique de santé et déclaration DRG à l'INMSS. Schéma didactique réalisé à partir de [321-323,465].

Le schéma montre pourquoi les erreurs administratives commises à l'hôpital ont des effets au-delà de ses murs. Un code de diagnostic incorrect dans la FOCG modifie le groupage DRG, donc le financement. Un CNP erroné à l'admission empêche la validation du service. Pour l'interne, le codage correct fait partie de l'acte médical.

12.4.3 La plateforme modernisée, un fait en mouvement

En octobre 2025, la CNAS annonçait que la PIAS serait remplacée "en 2026" par un système informatique moderne, financé par le Plan national de relance et de résilience (PNRR) [466]. Le communiqué final du 31 août 2026 emploie toutefois une autre formulation. Le projet "Redimensionarea, standardizarea și optimizarea Platformei Informatice din Asigurările de Sănătate - PIAS" (redimensionnement, standardisation et optimisation de la plateforme informatique de l'assurance maladie ; PNRR, composante C7, investissement I3.1), mené du 26 mars 2024 au 31 août 2026, a construit une nouvelle infrastructure technologique et y a migré la plateforme existante, avec ses données et ses services [323]. De nouveaux modules ont été ajoutés : le portail e-Sănătatea Mea, les formulaires médicaux électroniques, le Registre des assurés et les relations internationales. Le montant approuvé s'élevait à 493 900 milliers de lei hors TVA, dont 434 679 milliers de lei ont été payés [323]. Le document parle d'une "PIAS modernisée", et non d'un nouveau système portant un autre nom qui remplacerait le SIUI.

La situation en Roumanie en 2026

- Le 1er septembre 2026, la CNAS a lancé le portail eSănătateaMea, avec authentification par ROeID. L'assuré y voit son statut d'assuré, ses ordonnances des 5 dernières années, son historique de santé des 10 dernières années, les remboursements, l'état de sa carte, et peut prendre des rendez-vous en ligne [324].
- À cette même date, l'utilisation de la CEI n'est pas devenue obligatoire pour les prestataires, et l'accès des assurés aux services n'est pas conditionné par la CEI [325].

12.5 La transformation numérique des hôpitaux et la maturité numérique

12.5.1 Les investissements du PNRR

La numérisation de la santé est financée dans le PNRR au titre de la composante 7, "Transformation numérique", par l'investissement I3, consacré au système d'e-santé et à la télémédecine, et non au titre de la composante 12, qui finance surtout des bâtiments et des équipements [368]. Outre la modernisation de la PIAS (I3.1), l'investissement I3.2 a alloué 496 380 000 lei hors TVA à la numérisation de 13 institutions placées sous l'autorité du ministère de la Santé, comme les directions de santé publique et les services d'ambulance [467]. L'investissement I3.3 a alloué 493 900 000 lei hors TVA à des systèmes informatiques et à des infrastructures numériques dans 200 établissements de santé publics, avec un appel ouvert le 7 juillet 2023 et clos, après prolongation, le 15 septembre 2023 [326]. La date limite européenne pour atteindre les jalons et les cibles était le 31 août 2026 [468]. Au 27 septembre 2026, aucune synthèse officielle des résultats obtenus par les hôpitaux financés au titre de l'I3.3 n'avait été identifiée.

12.5.2 Comment mesurer la maturité numérique

Un hôpital qui a acheté des ordinateurs n'est pas, de ce seul fait, un hôpital numérique. Les modèles de maturité décrivent les étapes que franchit une organisation. Le plus connu, le modèle d'adoption du dossier médical électronique (*Electronic Medical Record Adoption Model*, EMRAM) de l'organisation HIMSS, comporte huit niveaux, de 0 (absence de systèmes départementaux) à 7 (dossier dynamique, analyse avancée des données), en passant par les systèmes auxiliaires, l'entrepôt de données cliniques, la documentation électronique, les prescriptions électroniques assorties d'une gouvernance, l'intégration des données et l'échange avancé de données [469]. Sur 1 026 hôpitaux des États-Unis, ceux des niveaux 6-7 avaient 3,25 fois plus de chances d'obtenir une meilleure note de sécurité que ceux du niveau 0 (IC à 95 % 2,33-4,55) [470]. L'étude est transversale ; elle montre donc une association, et non une relation causale.

L'EMRAM n'est pas le seul instrument. Une revue systématique a recensé 27 modèles de maturité numérique pour les hôpitaux et les a synthétisés en sept dimensions : stratégie, capacité technologique, interopérabilité, gouvernance, soins centrés sur le patient, personnes et compétences, analyse des données [471]. L'Allemagne offre une leçon utile pour la Roumanie. Le programme national DigitalRadar a évalué environ 90 % des hôpitaux allemands [449], et l'association entre le score de numérisation des processus et les indicateurs de qualité était en grande partie non significative sur le plan statistique ou difficile à interpréter [472]. Un panel de 49 hôpitaux allemands a montré que, jusqu'en 2025, des progrès nets n'étaient apparus que pour la

gestion des médicaments et pour la sortie, alors que de nombreux projets financés étaient encore en cours [473]. Le financement seul ne garantit pas le changement des façons de travailler.

12.6 Indicateurs et tableaux de bord

12.6.1 Les indicateurs d'activité

Les données des modules ADT et de codage permettent de calculer chaque mois quelques indicateurs standard. La durée moyenne de séjour (DMS) est le rapport entre le nombre total de journées d'hospitalisation et le nombre de sorties [474]. En 2023, dans la plupart des États de l'UE, la DMS des patients hospitalisés variait entre 5,2 et 9,4 jours [474]. Le taux d'occupation des lits se calcule en divisant les journées d'hospitalisation réalisées par les journées-lits disponibles (lits x jours calendaires), puis en multipliant par 100. En 2024, la Roumanie comptait 731 lits d'hôpital pour 100 000 habitants, le troisième niveau de l'UE, contre une moyenne européenne de 507 [475]. Un grand nombre de lits faiblement occupés indique une capacité utilisée de façon inefficace.

L'indice de case-mix (ICM) est la moyenne des poids relatifs des cas sortis au cours d'une période. Un ICM de 1,25 indique que l'hôpital a traité des cas dont la consommation de ressources attendue dépasse de 25 % celle du cas moyen de référence (le système DRG est expliqué au chapitre 4). Dans une base de données de 715 hôpitaux des États-Unis, l'ICM moyen est passé de 1,70 à la fin de 2019 à 1,79 au premier trimestre de 2024 [476].

Le taux de réhospitalisation à 30 jours est la proportion des sorties suivies d'une nouvelle hospitalisation dans les 30 jours. Il paraît simple, mais son interprétation demande de l'attention. Dans un service de neurochirurgie comptant 8 878 patients, ce taux a été de 2,9 %, et les auteurs ont conclu que l'indicateur n'a pas de définition homogène, qu'il dépend du système de santé et qu'il exige un ajustement sur le risque [477]. Une réhospitalisation programmée (par exemple, pour la deuxième cure de chimiothérapie) n'est pas un échec.

12.6.2 Le tableau de bord

Un tableau de bord (*dashboard*) affiche les indicateurs essentiels sur un seul écran, avec des seuils et une signalisation. La figure 12.4 en montre un exemple construit avec des données fictives. Une revue systématique de 54 articles a énuméré les exigences fonctionnelles des tableaux de bord hospitaliers (production de rapports, alertes, personnalisation, suivi des indicateurs de performance) et les exigences non fonctionnelles (rapidité, sécurité, intégration avec d'autres systèmes, entrepôt de données) [478]. Quant aux effets, une revue de 70 études a constaté que, sur les 43 résultats rapportés pour la durée de séjour, 28 indiquaient une réduction, alors que la mortalité restait inchangée dans 15 résultats sur 20 [479]. Les données d'activité recueillies à l'hôpital servent aussi à évaluer un service clinique. Par exemple, l'analyse rétrospective de 867 patients opérés d'un cancer du côlon à l'Hôpital clinique départemental d'urgence "Pius Brânzeu" de Timișoara, entre février 2019 et février 2023, a montré moins d'interventions programmées, davantage d'interventions en urgence et une durée moyenne de séjour plus courte pendant les années de la pandémie de COVID-19 que l'année précédente [115].

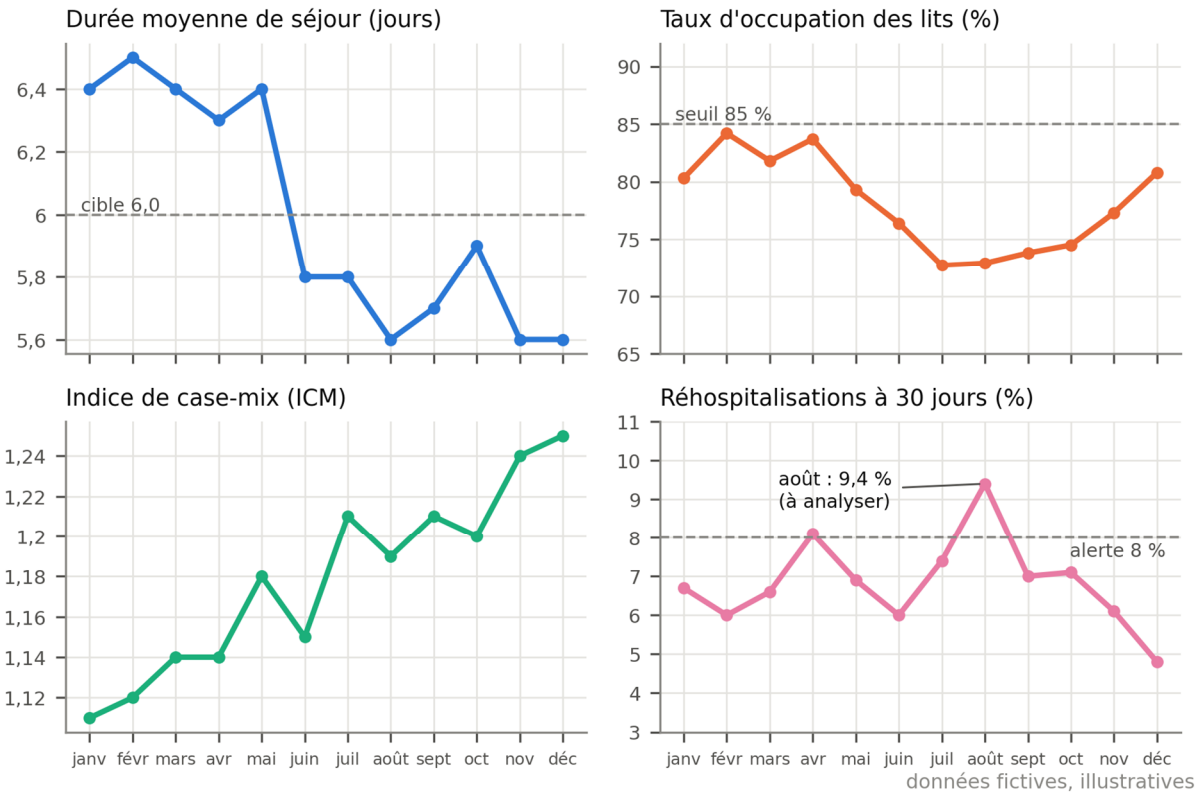


Figure 12.4. Tableau de bord didactique présentant quatre indicateurs mensuels d'un hôpital : la durée moyenne de séjour, le taux d'occupation des lits, l'indice de case-mix et le taux de réhospitalisation à 30 jours, avec des seuils de référence. Données fictives, illustratives.

Pour lire correctement le tableau de bord de la figure 12.4, il faut mettre les indicateurs en relation les uns avec les autres. La DMS diminue après mai, tandis que l'ICM augmente. L'hôpital traite des cas plus complexes en moins de jours, ce qui ressemble à un gain d'efficience. Le pic de réhospitalisations d'août (9,4 %) peut toutefois être le signe de sorties prématurées. Un bon tableau de bord ne donne pas la conclusion ; il montre où il faut la chercher.

12.7 La résilience : quand le système s'arrête

12.7.1 Les cyberattaques

Selon les communiqués de la Direction nationale de la cybersécurité (Directoratul Național de Securitate Cibernetică, DNSC), du 10 au 12 février 2024, le rançongiciel Backmydata, de la famille Phobos, a chiffré les données des serveurs de plusieurs hôpitaux roumains qui utilisaient le système d'information hospitalier Hipocrate. La DNSC a d'abord confirmé 21 hôpitaux touchés, et 79 autres établissements de santé ont été déconnectés d'Internet pour les besoins de l'enquête. Toujours selon la DNSC, la plupart des hôpitaux touchés disposaient de copies de sauvegarde datant d'un à trois jours, ce qui pouvait faciliter la restauration des données [186]. Le préambule de l'OUG n° 155/2024 mentionne un incident survenu au premier trimestre de 2024, qui "a touché 26 hôpitaux, au niveau national, par l'intermédiaire d'un fournisseur de services gérés" (notre traduction) [187]. Pour l'architecture des systèmes, la leçon est directe : un fournisseur commun à des dizaines d'hôpitaux devient un point unique de vulnérabilité. Le déroulement de l'attaque et le cadre juridique de la cybersécurité sont présentés au chapitre 5.

Le problème est international. Entre 2016 et 2021, 374 attaques par rançongiciel contre des prestataires de soins des États-Unis ont exposé les données de près de 42 millions de patients.

Leur nombre annuel a doublé, passant de 43 à 91, et 44,4 % des attaques ont perturbé les soins, le plus souvent par l'arrêt des systèmes électroniques [183]. Les effets dépassent l'hôpital attaqué. Pendant une attaque d'un mois contre quatre hôpitaux de San Diego, les deux services d'urgence d'un établissement voisin, non touché, ont accueilli en moyenne 251,4 patients par jour, contre 218,4 avant l'attaque [184]. En Irlande, l'attaque de mai 2021 a arrêté tous les systèmes informatiques d'un centre universitaire de cancérologie : la radiothérapie a été interrompue pour 113 patients, et les retards de compte rendu et de stockage des données n'ont été rattrapés qu'au 210e jour [480].

12.7.2 Le fonctionnement en mode dégradé

Les procédures de fonctionnement en mode dégradé (*downtime procedures*) décrivent comment l'hôpital travaille sans système. Winter recommande un plan de continuité d'activité qui indique comment les fonctions de l'hôpital sont assurées même en cas de panne majeure [448]. Un exemple publié de procédure complète comprend des exercices périodiques, un "visualiseur de secours" contenant les données essentielles des patients hospitalisés, des formulaires papier préparés à l'avance, le suivi de tous les changements survenus pendant la panne et une phase de retour à la normale (*uptime*) avec la conciliation médicamenteuse et la saisie rétroactive de la documentation [481]. Le cahier des charges SIEC-ATI exige lui aussi que le système de réanimation puisse créer temporairement l'enregistrement du patient lorsque le SIH de l'hôpital ne répond pas [244].

Pièges fréquents

- Confondre le système d'information (tout ce qui relève de l'information à l'hôpital, y compris le papier) avec le système informatique (la partie informatisée).
- Croire que relier directement chaque application à chacune des autres est plus simple qu'un moteur d'interfaces. Pour 10 applications, il faut 45 interfaces au lieu de 10.
- Interpréter un taux d'alertes ignorées de 90 % comme la preuve de la négligence des médecins. Le plus souvent, l'alerte est mal paramétrée.
- Lire la DMS isolément. Une DMS courte, accompagnée de nombreuses réhospitalisations, n'est pas un signe d'efficience.
- Supposer qu'un hôpital de niveau EMRAM élevé obtient automatiquement de meilleurs résultats cliniques. Les études disponibles montrent des associations, non des relations causales.
- Considérer que les procédures papier deviennent inutiles une fois l'hôpital informatisé.

À retenir

- Le SIH se compose de modules (ADT, FOCG, CPOE, pharmacie, LIS, RIS/PACS, bloc opératoire, réanimation, finances) qui doivent échanger des données de manière cohérente.
- Le moteur d'interfaces réduit le nombre de liaisons de n(n-1)/2 à n.
- Le CPOE a été associé à moins d'erreurs de prescription, mais il introduit de nouvelles erreurs, générées par la technologie.
- Les alertes d'interaction sont ignorées dans environ 90 % des cas. Leur optimisation revient à un comité pluridisciplinaire.
- Un bon cahier des charges contient des exigences mesurables et neutres vis-à-vis du fabricant.
- La DMS, le taux d'occupation, l'ICM et les réhospitalisations s'interprètent ensemble.
- Tout hôpital informatisé a besoin de procédures de fonctionnement en mode dégradé exercées périodiquement.

Questions d'autoévaluation

1. (QCS) Un hôpital dispose de 12 applications cliniques qui doivent échanger des données entre elles. Combien d'interfaces sont nécessaires si chaque application est reliée directement à chacune des autres ?

a) 12
b) 24
c) 66
d) 132
e) 144

2. (QCS) Quel module du système d'information hospitalier constitue la source de référence pour les données administratives du patient (identité, admission, transfert, sortie) ?

a) LIS
b) PACS
c) ADT
d) CPOE
e) l'entrepôt de données

3. (QCS) Au cours d'un mois de 30 jours, un service de 20 lits a réalisé 480 journées d'hospitalisation. Le taux d'occupation des lits est de :

a) 24 %
b) 48 %
c) 62,5 %
d) 80 %
e) 96 %

4. (QCM) À propos des alertes des systèmes d'aide à la décision clinique, les affirmations correctes sont :

a) une méta-analyse a estimé que les alertes d'interaction médicamenteuse sont ignorées dans environ 90 % des cas
b) la fatigue face aux alertes a une définition opérationnelle unique, utilisée par toutes les revues systématiques
c) les hôpitaux qui ont publié leur processus d'optimisation des alertes ont eu recours à un comité pluridisciplinaire
d) une alerte ignorée est toujours une erreur du médecin
e) le taux d'alertes ignorées et le taux d'acceptation sont des indicateurs fréquemment utilisés pour les alertes

5. (QCM) Lesquels des éléments suivants font partie d'une procédure de fonctionnement en mode dégradé bien conçue ?

a) des exercices périodiques de travail sans système
b) un accès de secours aux données essentielles des patients hospitalisés
c) la conciliation médicamenteuse après le redémarrage du système
d) l'élimination de la documentation papier après le retour à la normale, sans la saisir dans le système
e) des formulaires papier préparés à l'avance

CHAPITRE 13

Chapitre 13. L'intelligence artificielle en médecine : méthodes, évaluation, réglementation (2026)

Objectifs

À la fin de ce chapitre, l'étudiant sera capable :

- de distinguer un système fondé sur des règles d'un modèle obtenu par apprentissage automatique et de rattacher un exemple clinique à l'une des familles de méthodes (apprentissage supervisé, non supervisé, par renforcement, apprentissage profond, grands modèles de langage)
- de décrire les étapes du cycle de vie d'un modèle prédictif, de la question clinique à la surveillance après la mise en œuvre, et de reconnaître une fuite de données dans un protocole d'étude
- de calculer, à partir d'une matrice de confusion, la sensibilité, la spécificité, les valeurs prédictives, l'exactitude et le score F1, puis de recalculer les valeurs prédictives pour une autre prévalence
- d'interpréter une courbe ROC, l'aire sous la courbe et un graphique de calibration, sans confondre la discrimination avec l'utilité clinique
- de choisir la ligne directrice de rédaction adaptée (TRIPOD+AI, CONSORT-AI, SPIRIT-AI, DECIDE-AI, CLAIM, STARD-AI) pour un type d'étude donné
- d'énumérer les principaux risques des outils d'IA et les obligations qui découlent, en 2026, de l'AI Act, du MDR et des recommandations de l'ICMJE.

13.1 Qu'est-ce que l'intelligence artificielle en médecine ?

13.1.1 Définitions et les deux grandes traditions

Le terme d'intelligence artificielle (*artificial intelligence*, AI ; dans le texte, IA) n'a pas de définition unique. Une définition opérationnelle employée dans un manuel récent d'IA médicale décrit l'IA comme la science et la technologie des systèmes de calcul qui résolvent de manière autonome des problèmes inférentiels difficiles, et l'apprentissage automatique (*machine learning*, ML) comme la science et la technologie des systèmes qui apprennent à résoudre des problèmes en analysant des données liées à ces problèmes [482]. L'apprentissage automatique est donc un domaine de l'IA, et non un synonyme. Les mêmes auteurs avertissent que réduire toute l'IA à un seul sous-domaine, par exemple à l'apprentissage profond, empêche de choisir la méthode adaptée au problème clinique [482]. Pour une introduction en roumain, illustrée d'exemples tirés de la pratique médicale, on peut consulter [3], et, pour la place de l'IA dans l'ensemble de l'informatique médicale, [2].

La première tradition est l'IA **symbolique**. Les connaissances sont écrites explicitement par des experts, sous forme de règles et de relations logiques ("SI fièvre ET hémoculture positive

ALORS..."), et le système procède par déduction. À cette famille appartiennent les systèmes fondés sur des règles, avec des exemples historiques comme MYCIN, et l'aide à la décision par des règles indépendantes écrites en Arden Syntax [483]. Les alertes d'interactions médicamenteuses de la prescription électronique fonctionnent souvent de cette manière. L'avantage est la transparence : chaque règle peut être lue et vérifiée. L'inconvénient apparaît lorsque les connaissances doivent être entretenues manuellement pour des milliers de situations.

La seconde tradition est l'IA **fondée sur les données**. Ici, le modèle ne reçoit pas de règles, mais des exemples, à partir desquels il construit lui-même la relation entre les entrées et les sorties. Les réseaux de neurones, les modèles probabilistes et l'apprentissage par renforcement appartiennent à cette famille. Aliferis et Simon les qualifient de systèmes "superficiels sur le plan ontologique" : même un réseau profond, doté de nombreuses couches cachées, ne contient pas de représentation explicite de la maladie, mais seulement une relation apprise entre les entrées et le résultat [483]. La distinction compte pour l'interprétation, car un modèle fondé sur les données peut avoir de bonnes performances sans rien "savoir" de la physiopathologie.

13.1.2 Types d'apprentissage

La majorité des applications cliniques font appel à l'**apprentissage supervisé**. Le modèle reçoit des exemples étiquetés (radiographies marquées "pneumothorax" ou "normal", séjours marqués "a développé un sepsis" ou "non") et apprend à prédire l'étiquette pour de nouveaux cas [484]. La qualité de l'étiquette limite directement la qualité du modèle.

Dans l'**apprentissage non supervisé**, les données n'ont pas d'étiquettes. L'algorithme cherche des structures, par exemple des groupes de patients aux profils semblables ou des valeurs aberrantes [484]. Les groupes obtenus sont des hypothèses, non des diagnostics. Aliferis et Simon montrent que le regroupement non supervisé ne doit pas servir de substitut à un modèle prédictif, car un "bon" regroupement dépend du résultat qui nous intéresse, que l'algorithme ne voit pas [485].

L'**apprentissage par renforcement** (*reinforcement learning*) entraîne l'algorithme par essais et erreurs : les comportements qui rapprochent du résultat souhaité sont "récompensés" et deviennent plus probables [484]. En 2020, son impact en médecine clinique restait faible [484], et, dans la revue de portée (*scoping review*) des ECR d'IA publiés jusqu'en 2023, seuls deux des 27 systèmes qui ne travaillaient pas sur des images utilisaient l'apprentissage par renforcement [486].

13.1.3 Réseaux de neurones profonds

Un réseau de neurones artificiels est une succession de couches d'unités de calcul simples. Chaque unité combine ses entrées de façon pondérée et applique une fonction non linéaire. Le réseau est dit "profond" lorsqu'il possède de nombreuses couches cachées, et les poids sont ajustés pendant l'entraînement de sorte que l'erreur sur les exemples connus diminue [483,484]. Les réseaux convolutifs ont dominé l'analyse des images médicales (voir chapitre 7). Le prix de cette performance tient au grand nombre d'exemples étiquetés nécessaires et à l'opacité du modèle, discutée dans la section 13.5.

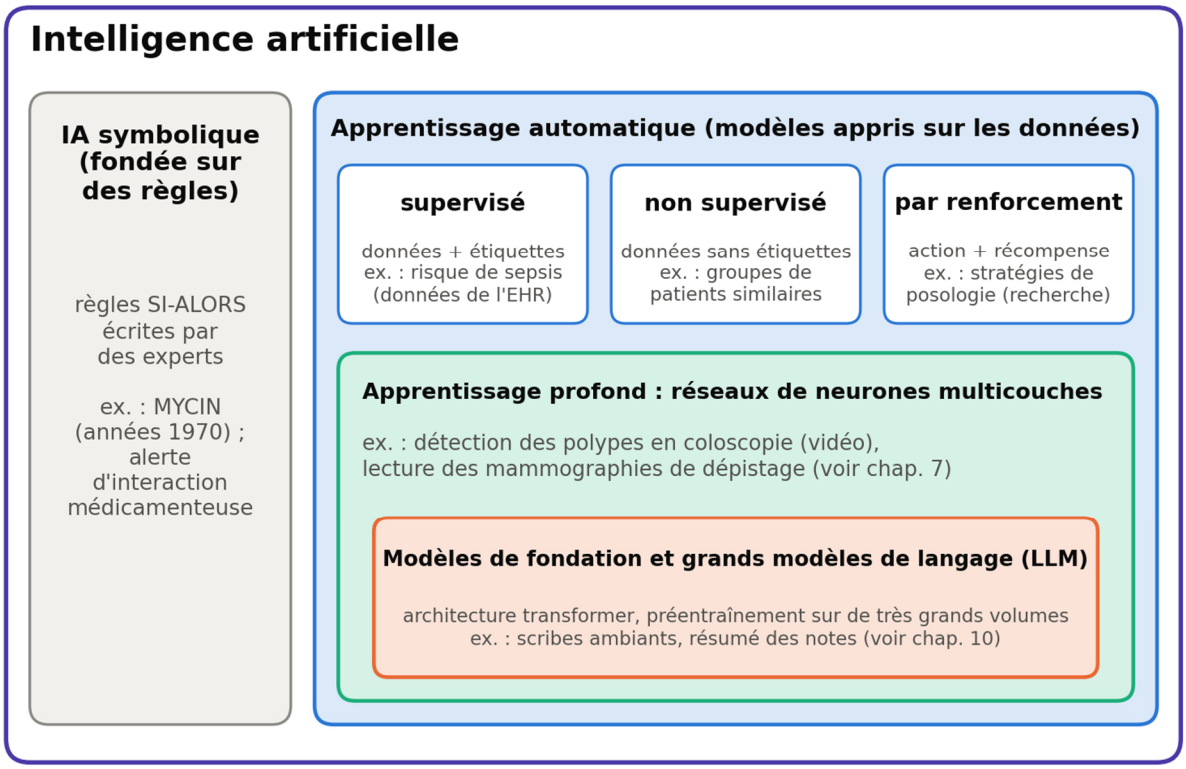


Figure 13.1. Les familles de méthodes d'intelligence artificielle et leurs relations, avec des exemples médicaux. L'apprentissage profond est une partie de l'apprentissage automatique, et les modèles de fondation et les grands modèles de langage sont, à leur tour, des réseaux profonds. Schéma didactique d'après [483,487].

13.1.4 Modèles de fondation et grands modèles de langage

Depuis 2017, l'architecture **transformer** est devenue la base de la majorité des modèles de traitement du langage. Elle utilise un mécanisme d'attention qui pondère, pour chaque mot, l'importance des autres mots du contexte [488]. Sur cette architecture ont été construits des modèles préentraînés sur de très grands volumes de texte, qui peuvent ensuite être adaptés à des tâches particulières par un ajustement fin (*fine-tuning*) beaucoup plus court que l'entraînement initial [488].

Un **modèle de fondation** (*foundation model*) est entraîné par autosupervision sur des jeux de données vastes et variés et peut accomplir de nombreuses tâches avec peu d'étiquettes spécifiques, voire aucune. Moor et ses collaborateurs ont proposé pour la médecine le concept d'IA médicale généraliste, capable de combiner images, données du dossier de santé électronique (DSE, *electronic health record*, EHR), résultats de laboratoire, génomique et texte [487]. Les **grands modèles de langage** (*large language models*, LLM) sont des modèles de fondation pour le texte. Ils répondent à des questions formulées librement sans entraînement dédié à cette tâche, et les chatbots (agents conversationnels) de type ChatGPT sont des LLM ajustés pour le dialogue [489].

La différence pratique avec un modèle classique tient au domaine d'application. Un score de risque pour une seule maladie est un système ciblé, dont les propriétés peuvent être vérifiées. Un LLM a un domaine illimité, ce qui rend difficile l'établissement de sa performance et de sa sécurité pour toutes les utilisations possibles [490]. Les applications des LLM en consultation et en mHealth sont traitées au chapitre 10.

13.2 Le cycle de vie d'un modèle

Un modèle utilisé dans les soins aux patients passe par des étapes distinctes, de la définition des exigences à l'intégration dans le flux de travail et à la surveillance [491]. La figure 13.2 les résume.

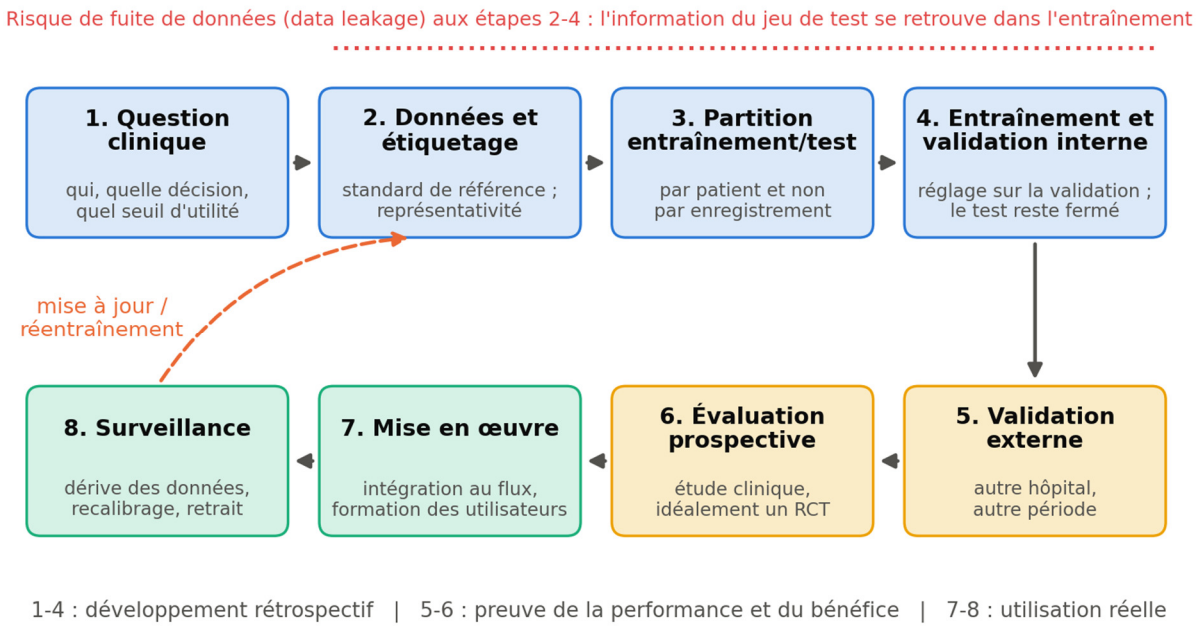


Figure 13.2. Le cycle de vie d'un modèle clinique d'IA : développement rétrospectif (1-4), preuve de la performance et du bénéfice (5-6), utilisation réelle (7-8), avec la boucle de mise à jour. La ligne pointillée marque la zone où apparaît habituellement la fuite de données. Schéma didactique d'après [491,492].

La question clinique. La première étape n'est pas le choix de l'algorithme, mais la définition de la décision que le modèle soutient : qui est le patient, à quel moment, quelle action suit un résultat positif et quelle performance minimale justifierait l'utilisation. Aliferis et Simon décrivent cette étape comme une ingénierie des exigences, avec des cibles de performance liées à des résultats cliniques concrets [491].

Les données et l'étiquetage. Le jeu de données doit ressembler à la population dans laquelle le modèle sera utilisé. L'étiquette est établie d'après un standard de référence (*reference standard*), terme que la ligne directrice CLAIM 2024 préfère à l'expression "ground truth" [493]. La variabilité de l'étiquetage et la sous-représentation de certains groupes de patients dans les données sont des sources de biais dès cette étape [494].

Entraînement, validation, test. Les données disponibles sont divisées en trois : un jeu d'entraînement (pour ajuster les poids du modèle), un jeu de validation (pour choisir la variante et les hyperparamètres) et un jeu de test, ouvert une seule fois, à la fin. La partition se fait par patient, et non par enregistrement ; sinon, le même patient apparaît dans les deux parties.

La validation externe. La performance obtenue sur les données propres (validation interne) est presque toujours optimiste. La validation externe utilise des données recueillies en un autre lieu ou à une autre période et explore l'hétérogénéité, l'équité et la généralisabilité [492]. Van Calster et ses collaborateurs ont formulé la conclusion dans leur titre : il n'existe pas de modèle "validé" une fois pour toutes, car les populations, les procédures de mesure et la pratique changent [495].

La mise en œuvre et la surveillance. Un modèle peut bien fonctionner dans une étude et mal dans un service. Au-delà de la limite du domaine qu'il a appris, la performance peut chuter brusquement, phénomène qu'Aliferis et Simon appellent la "falaise de connaissances" (*knowledge cliff*) [491]. La dérive du jeu de données (*dataset shift*) apparaît lorsque les données d'exploitation ne ressemblent plus à celles du développement, en raison de la technologie, de la population ou des comportements cliniques [496]. En avril 2020, l'hôpital de l'Université du Michigan a désactivé un modèle d'alerte au sepsis après que les changements liés à la COVID-19 avaient modifié la relation entre la fièvre et le sepsis bactérien [496]. Pour l'IA générative, le rapport de la NAM de 2025 recommande une surveillance continue de type "algorithmovigilance" [497].

13.2.1 La fuite de données

La fuite de données (*data leakage*) signifie que, pendant le développement, le modèle a accès à des informations dont il ne disposerait pas en utilisation réelle. Il en résulte une performance artificiellement gonflée. Kapoor et Narayanan ont trouvé des fuites de données dans 17 domaines scientifiques qui utilisent le ML, avec 294 articles concernés, et ont proposé une classification en huit types [498]. Exemples typiques en médecine :

- le même patient (ou la même lésion, photographiée deux fois) apparaît à la fois dans l'entraînement et dans le test
- la normalisation ou la sélection des variables est faite sur l'ensemble du jeu de données, avant la partition
- une variable contient, indirectement, le résultat : le code diagnostique de sepsis, saisi après confirmation, utilisé comme prédicteur du sepsis
- des données du futur servent à "prédire" le passé, par exemple des analyses prélevées après le moment où le modèle aurait dû donner l'alarme.

Lorsque les erreurs de ce type ont été corrigées dans une étude de reproduction (hors de la médecine, sur la prédiction des guerres civiles), les modèles complexes n'ont plus dépassé une simple régression logistique [498].

13.3 L'évaluation de la performance

13.3.1 La matrice de confusion et les indicateurs dérivés

Un classificateur binaire transforme un score continu (une probabilité) en une décision "positif/négatif" par le choix d'un seuil. Si l'on compare la décision au standard de référence, chaque patient tombe dans l'une des quatre cellules de la matrice de confusion : vrai positif (VP), faux positif (FP), faux négatif (FN), vrai négatif (VN). Les six premiers indicateurs du tableau 13.1 se calculent à partir de ces quatre nombres [499,500], et l'AUC découle des matrices obtenues pour tous les seuils possibles.

Tableau 13.1. Indicateurs de performance calculés à partir de la matrice de confusion, avec leur formule et leur interprétation. Source : [499-501].

Indicateur	Formule	Question à laquelle il répond	Dépend de la prévalence ?
Sensibilité (Se)	VP / (VP + FN)	Combien de malades sont détectés ?	Non (en principe)
Spécificité (Sp)	VN / (VN + FP)	Combien de sujets sains sont reconnus comme tels ?	Non (en principe)
Valeur prédictive positive (VPP)	VP / (VP + FP)	Si le test est positif, quelle est la probabilité d'être malade ?	Oui
Valeur prédictive négative (VPN)	VN / (VN + FN)	Si le test est négatif, quelle est la probabilité d'être sain ?	Oui
Exactitude	(VP + VN) / total	Quelle proportion des décisions est correcte ?	Oui
Score F1	2·VPP·Se / (VPP + Se)	Moyenne harmonique de la VPP et de la Se	Oui, ignore les VN
Aire sous la courbe ROC (AUC)	Probabilité qu'un malade ait un score plus élevé qu'un sujet sain	Dans quelle mesure le score sépare-t-il les deux groupes ?	Non

La sensibilité et la spécificité décrivent le test. Les valeurs prédictives décrivent la situation du patient qui a déjà un résultat, et elles varient avec la probabilité prétest, c'est-à-dire avec la prévalence dans la population testée [499]. L'exactitude peut tromper lorsque la maladie est rare : un "test" qui déclare tout le monde sain a une exactitude de 99 % dans une population où la prévalence est de 1 %, mais il ne détecte aucun malade.

13.3.2 La courbe ROC et l'aire sous la courbe

Chaque seuil produit un autre couple (Se, Sp). La courbe ROC (*receiver operating characteristic*) relie ces couples pour tous les seuils possibles, avec la sensibilité en ordonnée et 1 - spécificité en abscisse (figure 13.3). L'aire sous la courbe (angl. *area under the curve*, AUC) varie de 0,5, pour un test qui ne sépare pas mieux que le hasard, à 1, pour une séparation parfaite [501]. Les valeurs supérieures à 0,80 sont généralement considérées comme cliniquement utiles, mais la valeur ponctuelle s'interprète avec son intervalle de confiance à 95 % [501].

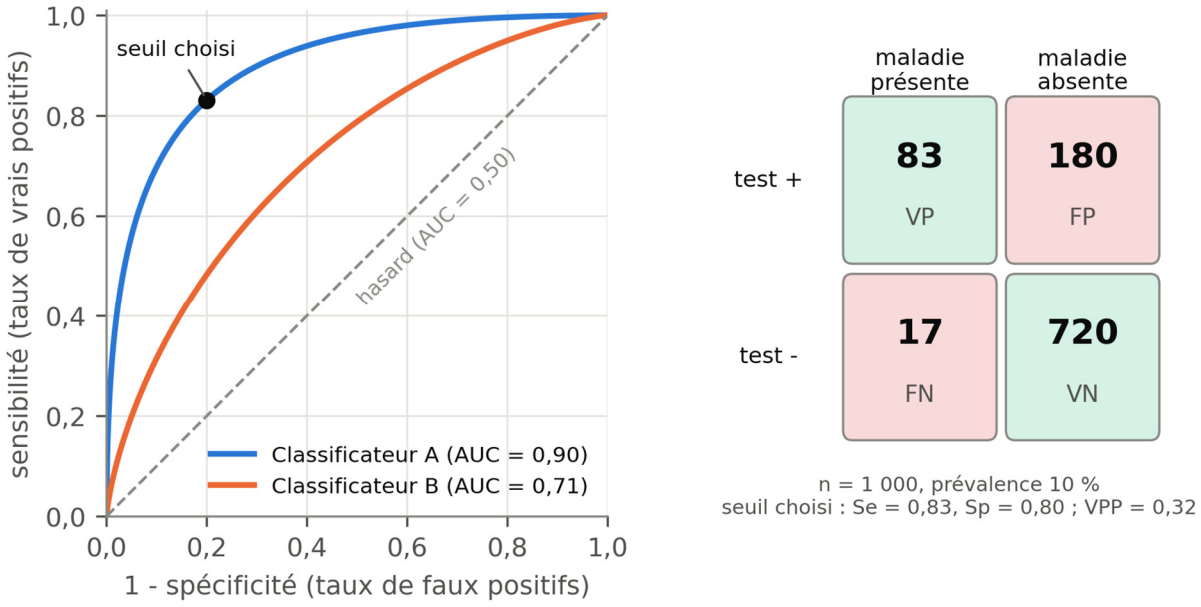


Figure 13.3. À gauche : les courbes ROC de deux classificateurs synthétiques, avec des scores distribués normalement chez les malades et chez les sujets sains (AUC 0,90 et 0,71). À droite : la matrice de confusion du classificateur A au seuil marqué (Se 0,83, Sp 0,80), pour 1 000 patients avec une prévalence de 10 %. Données synthétiques, illustratives, calculées dans le script de la figure.

Dans l'exemple de la figure 13.3, le classificateur A, au seuil choisi, détecte 83 malades sur 100. Pourtant, sur les 263 résultats positifs, 180 sont des faux positifs, si bien que la VPP n'est que de 83/263 = 0,32. Une bonne AUC ne garantit donc pas qu'un résultat positif soit fiable.

Un exemple réel montre pourquoi la validation externe compte. Le modèle de prédiction du sepsis d'Epic (Epic Sepsis Model, ESM), déployé dans des centaines d'hôpitaux aux États-Unis, a été évalué de manière externe sur 38 455 hospitalisations à Michigan Medicine. L'AUC était de 0,63 ; le modèle n'a pas identifié 67 % des patients atteints de sepsis et a généré des alertes pour 18 % de l'ensemble des patients hospitalisés [502]. La deuxième version du modèle, évaluée prospectivement dans quatre systèmes de santé sur 227 091 hospitalisations, a obtenu une AUC au niveau de l'hospitalisation comprise entre 0,82 et 0,92, mais la VPP est restée entre 0,13 et 0,26, avec une lourde charge d'alertes [503].

13.3.3 La dépendance à la prévalence : exemple numérique

Soit un outil d'IA avec Se = Sp = 0,90, appliqué à 10 000 personnes dans trois contextes (figure 13.4) :

- **dépistage dans la population générale, prévalence de 1 %** : 100 malades, dont 90 VP et 10 FN ; 9 900 sujets sains, dont 8 910 VN et 990 FP. VPP = 90/1 080 = 8,3 %. Sur 12 résultats positifs, 11 sont de fausses alarmes
- **patients adressés au spécialiste, prévalence de 10 %** : 900 VP, 100 FN, 900 FP, 8 100 VN. VPP = 900/1 800 = 50 %
- **patients symptomatiques du service spécialisé, prévalence de 30 %** : 2 700 VP, 300 FN, 700 FP, 6 300 VN. VPP = 2 700/3 400 = 79,4 %, tandis que la VPN descend à 6 300/6 600 = 95,5 %.

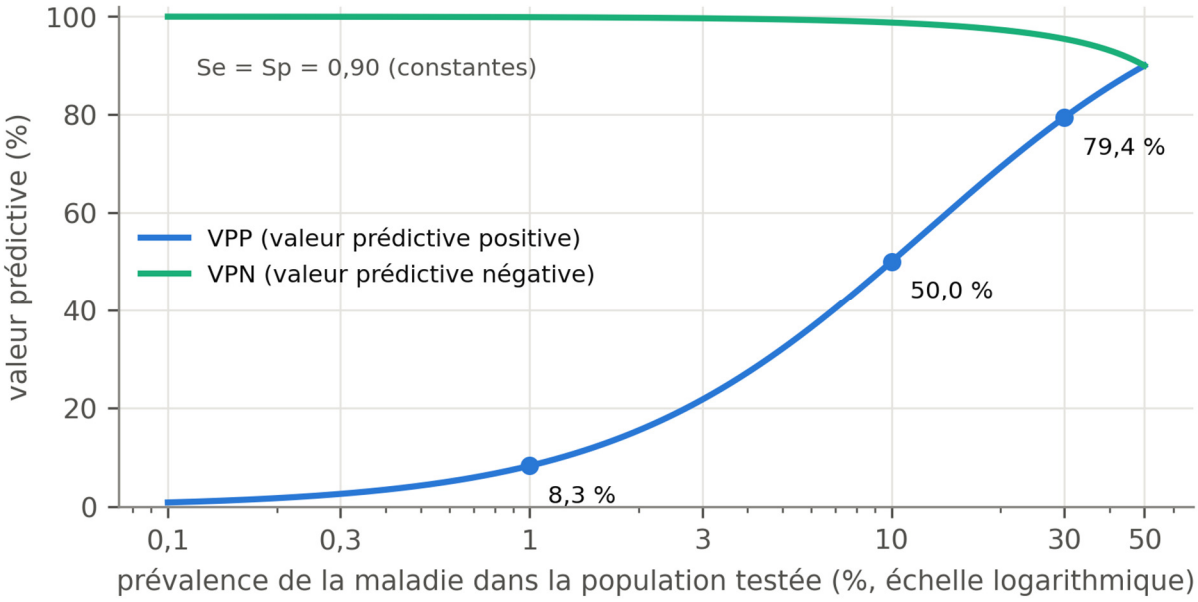


Figure 13.4. Les valeurs prédictives positive et négative en fonction de la prévalence, pour un test dont la sensibilité et la spécificité sont constantes et égales à 0,90. Les points marquent l'exemple numérique du texte (1 %, 10 %, 30 %). Calcul didactique, données synthétiques.

Le même algorithme, avec les mêmes Se et Sp, produit donc des résultats positifs de significations très différentes. C'est pourquoi la performance rapportée dans un centre tertiaire ne se transfère pas automatiquement aux soins primaires.

13.3.4 La calibration et l'utilité clinique

La discrimination (AUC) indique si le modèle ordonne correctement les patients. La **calibration** indique si les probabilités affichées sont justes : parmi les patients auxquels le modèle attribue un risque de 20 %, environ 20 % devraient présenter l'événement. Un modèle peut bien discriminer et néanmoins surestimer systématiquement le risque. Dans une analyse de 32 indicateurs de performance répartis en cinq domaines (discrimination, calibration, performance globale, classification, utilité clinique), Van Calster et ses collaborateurs recommandent de rapporter obligatoirement l'AUC, le graphique de calibration, une mesure d'utilité clinique, comme le bénéfice net issu de l'analyse de la courbe de décision, et la distribution des probabilités par catégorie de résultat [500]. Le score F1 n'a satisfait à aucun des deux critères de qualité analysés [500].

L'analyse de la courbe de décision (*decision curve analysis*) répond à une question simple : l'utilisation du modèle ferait-elle plus de bien que de mal ? Elle calcule le bénéfice net du modèle pour chaque probabilité seuil, c'est-à-dire pour le risque à partir duquel un clinicien déciderait d'intervenir [504]. Il est difficile de dire quelle doit être la valeur de l'AUC pour qu'un modèle mérite d'être utilisé, et la courbe de décision répond précisément à cette question [504].

13.3.5 Les lignes directrices de rédaction

Une étude d'IA ne peut être évaluée de façon critique que si elle est rapportée intégralement. Les lignes directrices du tableau 13.2 fixent les informations minimales, selon le type d'étude. TRIPOD+AI, publiée dans le BMJ en 2024, compte 27 éléments, s'applique aux modèles prédictifs, qu'ils soient construits par régression ou par ML, et remplace la liste TRIPOD de 2015 [505]. Le risque de biais de ces études s'évalue avec PROBAST+AI (2025) [506].

Tableau 13.2. Lignes directrices de rédaction pour les études qui évaluent des interventions et des modèles d'IA. Source : [493,505-511].

Ligne directrice	Année	Type d'étude	Contenu spécifique à l'IA
SPIRIT-AI	2020	Protocole d'essai clinique randomisé	15 éléments nouveaux par rapport à SPIRIT 2013
CONSORT-AI	2020	Rapport d'un essai clinique randomisé	14 éléments nouveaux par rapport à CONSORT 2010
DECIDE-AI	2022	Évaluation clinique précoce d'un système d'aide à la décision	17 éléments spécifiques à l'IA et 10 génériques
TRIPOD+AI	2024	Développement ou validation d'un modèle prédictif	27 éléments ; remplace TRIPOD 2015
CLAIM 2024	2024	Études d'IA en imagerie médicale	44 éléments, "standard de référence"
PROBAST+AI	2025	Évaluation du risque de biais (pas de rédaction)	16 + 18 questions de signalement
STARD-AI	2025	Exactitude diagnostique d'un test d'IA	18 éléments nouveaux ou modifiés
TRIPOD-LLM	2025	Études utilisant des grands modèles de langage	19 éléments principaux, 50 sous-éléments

CONSORT-AI exige, entre autres, la description de l'intervention d'IA, de la manière dont les données d'entrée et de sortie sont gérées, de l'interaction humain-IA, ainsi que l'analyse des cas d'erreur [507]. SPIRIT-AI a la même structure pour les protocoles [508]. DECIDE-AI couvre l'étape intermédiaire, au cours de laquelle un système est essayé à petite échelle en pratique, afin d'évaluer la sécurité et les facteurs humains avant une grande étude [509]. CLAIM 2024 recommande les termes "test interne" et "test externe" au lieu du mot ambigu "validation" [493].

13.4 Les preuves cliniques 2020-2026

La performance sur un jeu de test n'est pas la même chose que le bénéfice pour le patient. Une synthèse de 2022 a observé que les études prospectives commençaient à réduire l'écart entre la recherche et la mise en œuvre, surtout dans l'analyse des images, mais a aussi signalé des obstacles persistants, du manque de données aux biais [512]. Le bénéfice se démontre dans des études prospectives, idéalement dans des essais contrôlés randomisés (ECR ; *randomized controlled trials*, RCT). Une revue systématique publiée dans JAMA Network Open n'a trouvé, jusqu'en octobre 2021, que 41 ECR portant sur des interventions fondées sur le ML, avec une médiane de 294 participants. La moitié étaient monocentriques (51 %), 37 % concernaient l'endoscopie et aucun ne respectait toutes les exigences de CONSORT-AI [513].

Une revue de portée (*scoping review*) parue dans The Lancet Digital Health a actualisé le décompte jusqu'en novembre 2023 et a inclus 86 ECR dans lesquels l'IA était effectivement intégrée aux soins des patients [486]. La gastroentérologie représentait 43 % des études, suivie par la radiologie (13 %), et 63 % des études étaient monocentriques (figure 13.5). Sur les 86 études, 70 (81 %) ont rapporté un résultat favorable pour le critère principal. Les auteurs attirent toutefois l'attention sur le probable biais de publication, sur la pauvreté des données démographiques rapportées et sur le fait que seules 12 des 63 études publiées après 2021 citaient CONSORT-AI [486].

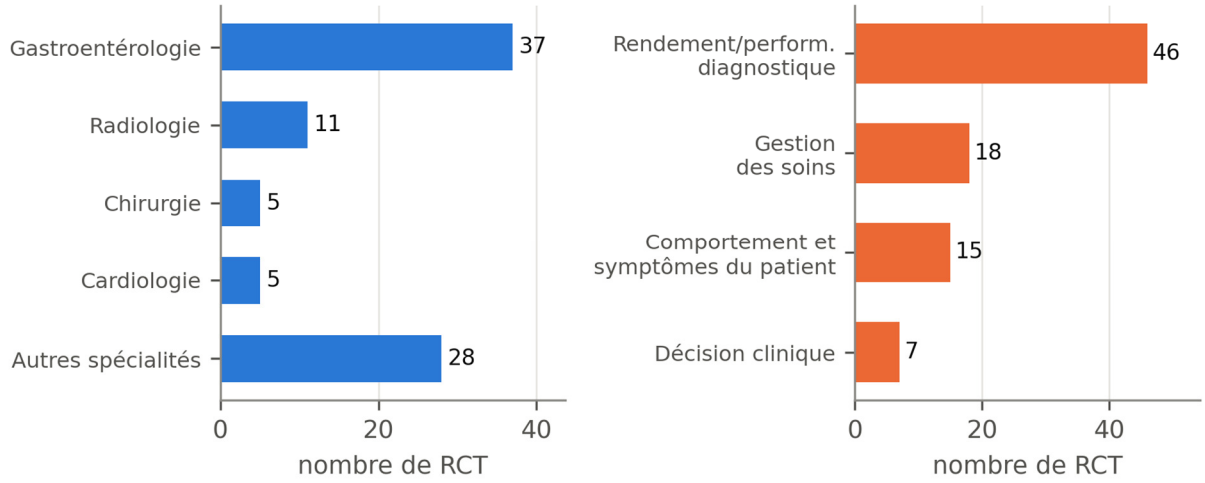


Figure 13.5. Les 86 ECR d'IA en pratique clinique publiés entre 2018 et novembre 2023 : à gauche, par spécialité ; à droite, par type de critère principal d'évaluation. Données réelles tirées de [486].

Quelques exemples représentatifs montrent à la fois ce qui a été obtenu et les limites :

- **La coloscopie assistée par l'IA.** Une méta-analyse de 28 ECR regroupant 23 861 participants a trouvé une augmentation de 20 % du taux de détection des adénomes (RR 1,20) et une diminution de 55 % du taux d'adénomes manqués, mais aucune différence pour les adénomes avancés et 39 % de résections supplémentaires de lésions non néoplasiques [514].
- **Le dépistage mammographique.** Dans l'étude suédoise MASAI, 105 934 femmes ont été randomisées entre une lecture assistée par l'IA et une double lecture standard. Le taux de cancers de l'intervalle s'est révélé non inférieur (1,55 contre 1,76 pour 1 000), la sensibilité est passée de 73,8 % à 80,5 %, et la spécificité est restée de 98,5 % dans les deux groupes [271]. Les détails techniques figurent au chapitre 7.
- **Le LLM comme aide au diagnostic.** Dans un ECR portant sur 50 médecins, l'accès à un LLM n'a pas amélioré le raisonnement diagnostique par rapport aux ressources conventionnelles (score médian de 76 % contre 74 %), bien que le LLM seul ait obtenu 16 points de pourcentage de plus que le groupe témoin [399]. La consultation assistée par un LLM est discutée au chapitre 10.
- **Les scribes ambiants.** Ces systèmes enregistrent la consultation et génèrent la note clinique. Dans un ECR pragmatique portant sur 238 médecins de 14 spécialités, l'une des deux applications testées a réduit de 9,5 % le temps consacré à la note par rapport au groupe témoin, l'autre n'a pas produit de modification statistiquement significative (-1,7 %), et les médecins ont signalé "occasionnellement" des inexactitudes cliniquement pertinentes [515]. Une étude randomisée en escalier (*stepped-wedge*) portant sur 66 cliniciens a trouvé une diminution de l'épuisement professionnel (-0,44 point sur une échelle de 5) et du temps consacré aux notes (-0,36 heure par jour), sans augmentation statistiquement significative de la satisfaction professionnelle [516]. Une troisième étude, croisée, portant sur 160 cliniciens, a comparé deux produits : tous deux ont réduit les scores d'épuisement, mais le temps de travail à domicile, après les heures de service, est resté en grande partie inchangé [517]. Le rapport de la NAM de 2025 range les scribes ambiants parmi les applications à risque plus faible, proches de l'utilisation courante [497].

Le tableau d'ensemble est donc contrasté. Les bénéfices sont le mieux documentés pour des tâches de détection en imagerie et en endoscopie, avec des critères intermédiaires (taux de détection) plus souvent que des critères cliniques finaux (mortalité, complications).

13.5 L'IA explicable

Beaucoup de modèles de ML sont des "boîtes noires" : ils produisent un score sans justification lisible. L'IA explicable (*explainable AI*, XAI) cherche à montrer ce qui a déterminé une prédiction. Deux familles de méthodes sont les plus fréquentes : l'attribution d'une importance à chaque variable et les cartes de saillance pour les images.

Les valeurs SHAP (*SHapley Additive exPlanations*) empruntent à la théorie des jeux l'idée de Lloyd Shapley d'un partage équitable du gain entre les participants à une coalition. Appliquée à un modèle, la méthode attribue à chaque variable une contribution à la prédiction pour un patient ("l'âge a augmenté le risque de x, le lactate de y"). Aliferis et Simon énumèrent toutefois neuf limites des valeurs de Shapley dans l'explication des modèles de ML, notamment le fait qu'elles n'expliquent même pas correctement les modèles linéaires et qu'elles peuvent attribuer de l'importance à des variables sans lien avec le résultat ; c'est pourquoi ils ne les recommandent pas comme méthode de routine [491].

Les cartes de saillance colorent les pixels qui ont le plus influencé la décision d'un réseau sur une image. Évaluées sur deux jeux publics de radiographies, les huit méthodes de saillance testées ont toutes échoué à au moins un critère de fiabilité (localisation, sensibilité à la randomisation des poids, répétabilité, reproductibilité) et ont localisé les lésions moins bien que les réseaux construits spécialement pour la segmentation [518].

Ghassemi et ses collaborateurs ont qualifié ces espoirs de "faux espoir" : les méthodes actuelles d'explication ne peuvent garantir, au niveau du patient individuel, que l'explication reflète la raison réelle de la prédiction. Ils recommandent une validation interne et externe rigoureuse comme moyen plus direct d'obtenir la confiance et mettent en garde contre l'imposition de l'explicabilité comme exigence pour les modèles cliniques [519]. Un exemple illustre le problème. Des réseaux profonds standard ont pu prédire l'origine ethnique déclarée du patient à partir de radiographies, d'examens TDM et de mammographies (AUC 0,81-0,99), alors qu'aucun expert humain ne voit cette information sur l'image [520]. Une carte de saillance n'aurait pas révélé un tel mécanisme caché.

13.6 Les risques

Biais et équité. Un modèle apprend à partir de données historiques et peut en reproduire les inégalités. Des classificateurs de radiographies thoraciques entraînés sur trois grands jeux de données ont sous-diagnostiqué de manière sélective des groupes défavorisés (femmes, patients noirs, patients de statut socio-économique faible), avec des taux plus élevés à l'intersection de ces catégories, par exemple chez les femmes hispaniques [194]. Les sources de biais apparaissent lors de l'acquisition des données, dans la variation génétique entre les populations et dans la variabilité de l'étiquetage [494]. Le code de conduite de la NAM fait de l'équité l'un de ses six engagements [521].

Le biais d'automatisation (*automation bias*). La tendance à suivre la suggestion d'un système automatique touche aussi les experts. Dans une expérience prospective, 27 radiologues ont lu 50 mammographies avec une prétendue assistance par l'IA. Lorsque la suggestion BI-RADS était erronée, la proportion de lectures correctes est tombée d'environ 80 % à 20 % (débutants), 25 % (expérience moyenne) et 46 % (radiologues très expérimentés) [281].

Hallucinations. Pour l'IA générative, une "hallucination" (ou confabulation) est une information incorrecte ou infondée, présentée de manière plausible, qui découle de la façon dont le modèle génère le texte par prédiction [497]. Lorsqu'on leur a demandé de retrouver la littérature

pertinente pour 11 revues systématiques, GPT-3.5, GPT-4 et Bard ont produit des références hallucinées (avec au moins deux éléments erronés parmi le titre, le premier auteur et l'année) dans respectivement 39,6 %, 28,6 % et 91,4 % des cas [44]. Toute référence proposée par un LLM se vérifie dans PubMed ou CrossRef avant d'être utilisée.

Sécurité. Les attaques ne visent pas seulement les serveurs. Le remplacement de seulement 0,001 % des unités de texte d'entraînement par de la désinformation médicale a produit des modèles qui propageaient des erreurs médicales, alors qu'ils obtenaient les mêmes scores aux tests usuels [522].

Responsabilité. Un outil d'IA ne peut pas être tenu pour responsable sur le plan juridique ou déontologique. Le principe de responsabilité du code de la NAM exige des actions identifiables et mesurables, une documentation claire des bénéfices et des contrôles clairs pour les éventuelles conséquences indésirables [521]. La décision et la signature restent celles du médecin.

Pièges fréquents

- Saisir les données d'un patient réel (nom, code numérique personnel ou CNP, images, antécédents) dans un chatbot public. Les données concernant la santé constituent une catégorie particulière, dont le traitement est interdit en règle générale et permis seulement dans les exceptions de l'article 9 du RGPD [118]. Un chatbot public n'est pas un outil approuvé par l'hôpital pour les traiter.
- Déléguer la décision : "l'IA l'a dit" n'est pas une justification clinique. La suggestion se vérifie, et le médecin répond de la décision.
- Présenter un texte généré comme son propre texte, dans un exposé, un mémoire de licence ou un article, sans le déclarer.
- Citer une référence proposée par un LLM sans la vérifier dans PubMed ou CrossRef.
- Juger un modèle uniquement sur l'AUC ou l'exactitude, sans tenir compte de la prévalence, de la calibration et de la validation externe.
- Confondre "validé en interne" et "dont le bénéfice pour les patients est démontré".

13.7 La réglementation en 2026

13.7.1 Le règlement européen sur l'IA (AI Act)

Le règlement (UE) 2024/1689, connu sous le nom d'AI Act, a été adopté le 13 juin 2024 et est entré en vigueur le 1er août 2024 [22]. En tant que règlement, il s'applique directement dans tous les États membres, y compris en Roumanie, sans loi de transposition. L'approche est fondée sur le risque, avec quatre niveaux (tableau 13.3) [523].

Tableau 13.3. Les niveaux de risque de l'AI Act, avec des exemples pertinents pour la santé. Source : [22,523].

Niveau	Signification	Exemples	Conséquence
Risque inacceptable	Pratiques interdites (art. 5)	Manipulation préjudiciable, notation sociale	Interdites depuis le 2 février 2025
Haut risque	Risque grave pour la santé, la sécurité ou les droits	Logiciel médical intégrant de l'IA, évalué par un organisme notifié, triage des patients aux urgences	Système de gestion des risques, données de qualité, documentation, contrôle humain
Risque de transparence	L'utilisateur doit savoir qu'il interagit avec une IA	Chatbots, contenu généré	Information de l'utilisateur, marquage du contenu synthétique
Risque minimal	La majorité des applications	Filtres antispam, jeux vidéo	Aucune obligation nouvelle

Le lien avec les dispositifs médicaux est direct. L'article 6, paragraphe 1, qualifie de système d'IA à haut risque un système d'IA qui est lui-même un produit, ou le composant de sécurité d'un produit, couvert par la législation de l'annexe I et soumis à une évaluation de la conformité par un tiers ; or le MDR et l'IVDR font partie de cette annexe [22]. En pratique, un logiciel médical intégrant de l'IA, de classe IIa ou supérieure, évalué par un organisme notifié, relève du régime des systèmes à haut risque. L'annexe III ajoute, entre autres, les systèmes de triage des patients aux urgences et l'IA utilisée pour l'évaluation des risques dans l'assurance maladie [22].

Le calendrier a été modifié en 2026. L'omnibus numérique sur l'IA, le règlement (UE) 2026/1744 du 8 juillet 2026, publié le 24 juillet 2026, a reporté les obligations applicables aux systèmes à haut risque [197]. Au 27 septembre 2026, la situation est la suivante :

- **2 février 2025** : s'appliquent les interdictions (art. 5) et l'obligation relative à la maîtrise de l'IA (art. 4) [22]. L'omnibus a reformulé l'article 4 : les fournisseurs et les déployeurs prennent des mesures pour favoriser la maîtrise de l'IA de leur personnel, sans avoir à garantir un niveau déterminé pour chaque personne [197].
- **2 août 2025** : s'appliquent les règles relatives aux modèles d'IA à usage général, la gouvernance et les sanctions [22].
- **2 décembre 2026** : échéance pour le marquage du contenu synthétique par les systèmes génératifs mis sur le marché avant le 2 août 2026 [197].
- **2 décembre 2027** : obligations applicables aux systèmes à haut risque autonomes, relevant de l'annexe III [197].
- **2 août 2028** : obligations applicables aux systèmes à haut risque intégrés dans des produits relevant de l'annexe I, donc aussi aux dispositifs médicaux intégrant de l'IA [197].

Par ailleurs, la Commission a proposé en décembre 2025 de simplifier le MDR et l'IVDR, notamment en les déplaçant dans la section B de l'annexe I, ce qui limiterait l'application de l'AI Act aux dispositifs médicaux [524]. Au 27 septembre 2026, il ne s'agit que d'une proposition.

13.7.2 Le MDR et les logiciels médicaux

Le règlement (UE) 2017/745 relatif aux dispositifs médicaux (MDR) s'applique depuis le 26 mai 2020 [284]. Pour les logiciels, la règle 11 de l'annexe VIII fixe la classe : un logiciel qui fournit des informations utilisées pour prendre des décisions à des fins diagnostiques ou thérapeutiques relève au moins de la classe IIa, de la classe IIb si la décision peut entraîner une détérioration grave de l'état de santé, et de la classe III si elle peut entraîner le décès ou une détérioration irréversible [284]. La qualification et la classification des logiciels sont expliquées dans le guide MDCG 2019-11, révisé en juin 2025 [372]. Un algorithme destiné par son fabricant à proposer des diagnostics ou des traitements est donc en général un dispositif médical, et non un simple "programme".

La protection des données demeure un cadre distinct. Les données concernant la santé constituent une catégorie particulière dans le RGPD [118], et la loi roumaine 190/2018 (Legea 190/2018) ne permet d'utiliser des données de santé pour des décisions automatisées ou pour le profilage qu'avec un consentement explicite ou en vertu de dispositions légales expresses, assorties de mesures de protection [176].

13.7.3 La FDA, les codes de conduite et le guide de l'OMS

Aux États-Unis, la FDA publie une liste des dispositifs médicaux qui utilisent l'IA. Le fichier officiel téléchargé depuis la page mise à jour le 4 septembre 2026 contenait 1 614 dispositifs, avec

des décisions jusqu'au 29 juin 2026, dont 1 230 (76,2 %) en radiologie, et 335 dispositifs ont été autorisés pour la seule année 2025 [267]. La FDA précise que la liste n'est pas exhaustive. L'évolution annuelle est présentée au chap. 7.

À côté de la législation existent des codes volontaires. Le code de conduite de l'Académie nationale de médecine des États-Unis (NAM, 2025) formule dix principes (engagé, sûr, efficace, équitable, efficient, accessible, transparent, responsable, sécurisé, adaptatif) et six engagements, parmi lesquels la protection de la santé humaine comme objectif principal, l'équité et la surveillance ouverte de la performance [521]. Le rapport de la NAM sur l'IA générative, paru la même année, classe les applications des LLM selon leur maturité : à court terme, l'éducation du patient et la synthèse de l'information ; à long terme, les assistants virtuels de santé [497]. L'OMS a publié le 18 janvier 2024 son guide d'éthique et de gouvernance pour les grands modèles multimodaux, qui comprend plus de 40 recommandations et cinq grands usages : diagnostic et soins cliniques, utilisation guidée par le patient, tâches administratives, formation médicale et infirmière, recherche scientifique et développement de médicaments [410,411].

La situation en Roumanie en 2026

- L'AI Act et le MDR s'appliquent directement, en tant que règlements européens. Au 27 septembre 2026, les obligations applicables aux dispositifs médicaux intégrant de l'IA en tant que systèmes à haut risque sont prévues pour le 2 août 2028 [197].
- Chaque État membre doit disposer d'au moins un bac à sable réglementaire de l'IA (*regulatory sandbox*) opérationnel au plus tard le 2 août 2027 [197].
- L'obligation de maîtrise de l'IA (art. 4) s'applique déjà aux hôpitaux qui déploient des systèmes d'IA, dans sa forme modifiée en juillet 2026 [197].
- Les données de santé utilisées pour des décisions automatisées relèvent de l'article 3 de la loi 190/2018 [176].
- Pour l'étudiant et l'interne, la conséquence pratique est simple : les outils publics d'IA ne reçoivent pas de données de patients, et tout résultat est vérifié et assumé par un être humain.

13.8 Note de renvoi

Les méthodes de recherche avancées (jumeaux numériques, apprentissage fédéré, application de TRIPOD+AI au développement proprement dit d'un modèle) dépassent le niveau de ce cours et sont traitées dans la monographie doctorale de l'auteur ; une synthèse de l'état des jumeaux numériques en oncologie se trouve dans [525].

Exemple

Ce qu'un outil d'IA peut et ne peut pas faire dans un manuscrit. Les recommandations de l'ICMJE, mises à jour en janvier 2026, fixent les règles [202] :

- Un chatbot ne peut pas être auteur, car il ne peut pas répondre de l'exactitude, de l'intégrité et de l'originalité du travail.
- L'utilisation de l'IA se déclare lors de la soumission, dans la lettre d'accompagnement et dans le manuscrit : l'aide à la rédaction, dans la section des remerciements ; l'utilisation pour la collecte des données, l'analyse ou la génération des figures, dans la section des méthodes.
- Le matériel généré par l'IA ne peut pas être cité comme source primaire. Les auteurs vérifient le texte, car il peut être incorrect, incomplet ou biaisé, et garantissent qu'il ne contient pas de plagiat, ni dans le texte ni dans les images.
- L'absence de déclaration peut imposer des corrections et, dans certaines situations, être considérée comme une faute.

Liste de vérification avant la soumission : (1) j'ai noté quel outil, quelle version et pour quoi je l'ai utilisé ; (2) j'ai vérifié chaque affirmation et chaque référence dans la source originale ; (3) je n'ai pas saisi de données de patients dans un service public ; (4) la déclaration figure dans la section appropriée ; (5) tous les auteurs humains assument le texte final.

À retenir

- L'apprentissage automatique est une partie de l'IA ; l'apprentissage profond et les LLM sont, à leur tour, des cas particuliers d'apprentissage automatique.
- Un modèle se juge sur des données externes, de manière prospective, et se surveille après sa mise en œuvre, car les données changent.
- La fuite de données gonfle artificiellement la performance. La partition se fait par patient, avant tout traitement.
- La sensibilité et la spécificité décrivent le test. La VPP et la VPN dépendent de la prévalence.
- L'AUC mesure la discrimination, et non la calibration ni l'utilité clinique.
- Un inventaire arrêté en novembre 2023 a trouvé 86 ECR dans lesquels l'IA était intégrée à la pratique, pour la plupart monocentriques et portant surtout sur des critères intermédiaires.
- Les explications post hoc (SHAP, cartes de saillance) ne remplacent pas la validation.
- L'IA ne peut pas être auteur et ne peut pas signer une décision. Son utilisation se déclare.

Questions d'autoévaluation

1. (QCS) Un outil d'IA a une sensibilité de 0,90 et une spécificité de 0,90. Appliqué dans une population où la prévalence de la maladie est de 1 %, sa valeur prédictive positive est d'environ :

a) 90 %
b) 50 %
c) 8 %
d) 1 %
e) 99 %

2. (QCS) Laquelle des situations suivantes constitue une fuite de données (*data leakage*) ?

a) l'utilisation d'un jeu de test provenant d'un autre hôpital
b) la présence des mêmes patients à la fois dans le jeu d'entraînement et dans le jeu de test
c) la communication de l'intervalle de confiance de l'AUC
d) le recalibrage du modèle après sa mise en œuvre
e) l'étiquetage des images par deux radiologues indépendants

3. (QCS) La ligne directrice de rédaction destinée aux études qui développent ou valident un modèle prédictif clinique, construit par régression ou par apprentissage automatique, est :

a) CONSORT-AI
b) SPIRIT-AI
c) DECIDE-AI
d) TRIPOD+AI
e) CLAIM 2024

4. (QCM) Concernant la courbe ROC et l'aire sous la courbe (AUC), sont vraies les affirmations suivantes :

a) la courbe ROC représente la sensibilité en fonction de 1 - spécificité, pour tous les seuils
b) une AUC de 0,5 correspond à une discrimination qui n'est pas meilleure que le hasard
c) une AUC élevée garantit que les probabilités affichées par le modèle sont bien calibrées
d) l'AUC ne dépend pas du seuil de décision choisi
e) une AUC de 0,90 garantit une VPP d'au moins 0,90

5. (QCM) Selon les recommandations de l'ICMJE relatives à l'utilisation de l'IA par les auteurs :

a) un chatbot peut figurer parmi les auteurs s'il a écrit la plus grande partie du texte
b) l'utilisation de l'IA pour la rédaction se décrit dans la section des remerciements
c) l'utilisation de l'IA pour l'analyse des données ou la génération des figures se décrit dans la section des méthodes
d) les auteurs humains répondent du contenu généré par l'IA
e) un texte généré par l'IA peut être cité comme source primaire

CHAPITRE 14

Chapitre 14. Synthèse : les compétences du médecin dans l'environnement numérique et l'éthique de l'utilisation de l'intelligence artificielle

Objectifs

À la fin de ce chapitre, l'étudiant pourra :

- placer les 13 chapitres précédents sur une carte à cinq couches (données, standards, systèmes, utilisation clinique, gouvernance) et expliquer, sur un exemple clinique, les liens qui les unissent
- vérifier, à l'aide d'une liste concrète, ses compétences numériques à la fin de la première année et indiquer le TP dans lequel il a exercé chacune d'elles
- situer correctement dans le temps les échéances de l'AI Act, du règlement (UE) 2026/1744, de l'EHDS, de NIS2 et d'eIDAS 2 qui modifieront la pratique médicale d'ici 2031
- appliquer une liste de décision avant d'utiliser un outil d'intelligence artificielle (IA) pour une tâche d'étude ou de travail clinique
- répondre à des questions de révision couvrant les chapitres 1-13

14.1 La carte de la discipline

Un étudiant qui a parcouru 13 chapitres peut avoir l'impression d'avoir appris 13 sujets distincts : théorie de l'information, bases de données, codes, RGPD, signaux, images, dossiers, télémédecine, applications mobiles, signatures, hôpitaux, intelligence artificielle. Le lien entre eux devient visible si l'on suit le parcours d'une seule information clinique. Une pression artérielle de 150/95 mmHg, mesurée chez un patient fictif, est d'abord une donnée à cinq composantes : patient, paramètre, valeur, moment, méthode (chapitre 2). Pour être conservée, elle entre dans une table dotée d'une clé primaire et de règles de validation (chapitre 3). Pour être comprise par un autre système, elle reçoit un code LOINC et devient une ressource FHIR Observation à deux composantes, la systolique et la diastolique (chapitres 4 et 8, TP 8). Elle arrive ensuite dans le dossier de santé électronique (DSE, *electronic health record*, EHR) de l'hôpital, sur la plateforme de la caisse d'assurance maladie ou sur le portail du patient (chapitres 11 et 12). Le médecin l'utilise lors d'une téléconsultation ou dans un score de risque (chapitres 9, 10 et 13). Tout au long de ce parcours, les règles de confidentialité, de sécurité et de responsabilité déterminent qui la voit et ce que l'on peut en faire (chapitres 5 et 11).

La figure 14.1 résume ce parcours en cinq couches. L'ordre n'est pas fortuit. Chaque couche suppose que les couches inférieures sont maîtrisées : un algorithme d'IA entraîné sur des données non codées ou incomplètes sera médiocre, aussi sophistiqué soit-il, et une règle juridique bien rédigée ne protège rien si l'hôpital n'a ni copies de sauvegarde ni journaux d'audit. La boucle en pointillé de la figure montre que la gouvernance agit en retour sur les données. Le RGPD, l'EHDS

et l'AI Act précisent quelles données peuvent être collectées, combien de temps elles sont conservées et à quelles conditions elles peuvent être réutilisées pour la recherche.

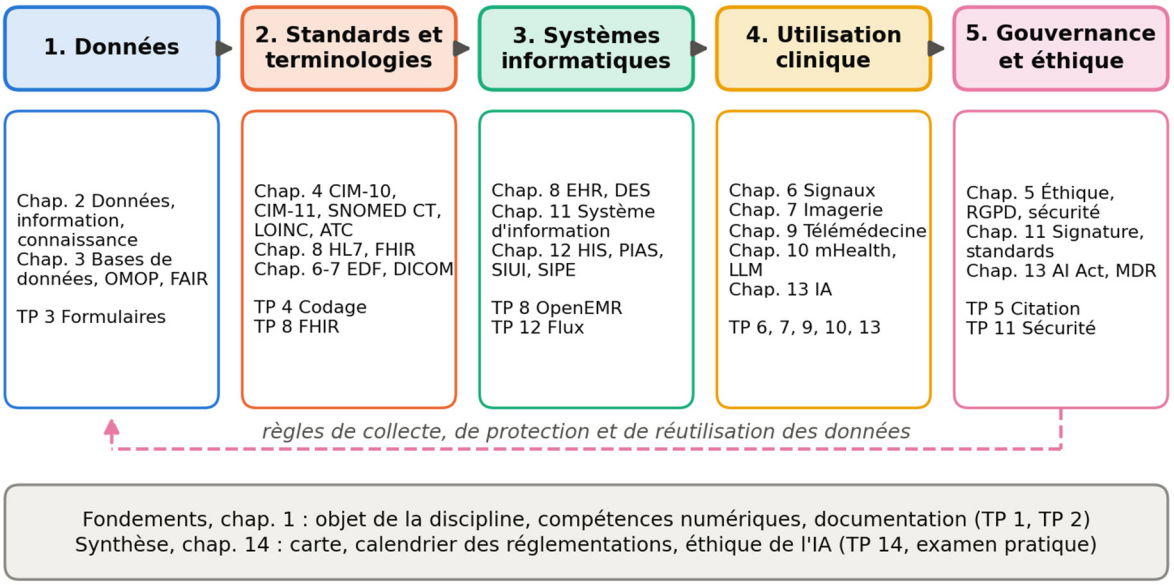


Figure 14.1. La carte de la discipline : les cinq couches (données, standards et terminologies, systèmes informatiques, utilisation clinique, gouvernance et éthique), avec les chapitres et les travaux pratiques de l'ouvrage. La boucle en pointillé montre l'effet des règles sur la collecte et la réutilisation des données. Schéma didactique.

La carte aide aussi à lire la littérature. Lorsqu'un article annonce qu'un modèle d'IA "a surpassé les médecins", les questions se posent couche par couche : quelles données ont été utilisées et dans quelle mesure elles étaient représentatives, quel standard de référence a établi l'étiquette, dans quel système le modèle a été intégré, quel résultat clinique a été mesuré et qui répond en cas d'erreur. Le chapitre 13 a montré qu'un inventaire mis à jour jusqu'en novembre 2023 a recensé 86 essais randomisés dans lesquels l'IA était effectivement intégrée aux soins, pour la plupart monocentriques [486]. La carte explique pourquoi : chaque couche ajoute des conditions qu'une étude rétrospective ne vérifie pas.

14.2 Les compétences à la fin de la première année

Le chapitre 1 a présenté trois cadres de référence. Les recommandations de l'Association internationale d'informatique médicale (IMIA), révisées en 2023 à partir d'une étude Delphi internationale, décrivent les besoins de formation de trois rôles, et, pour l'étudiant en médecine, le rôle pertinent est celui d'utilisateur compétent [29]. Le cadre DECODE regroupe les compétences en santé numérique de la formation médicale en quatre domaines : le professionnalisme en santé numérique, la santé numérique du patient et de la population, les systèmes d'information de santé et la science des données de santé [30]. Le cadre européen DigComp 3.0 conserve cinq domaines de compétence et intègre dans chacun d'eux les aspects liés à l'IA [34].

Pour les professionnels de santé de l'UE, le Policy Brief 71 de l'Observatoire européen des systèmes et des politiques de santé a constaté que l'accent portait sur la maîtrise de l'information et des données, la communication et la collaboration, la sûreté et la sécurité, la résolution de problèmes et la création de contenus numériques. Le même document juge l'offre de formation insuffisante et note que les participants roumains ont signalé l'obstacle de la langue anglaise [35]. L'édition roumaine de ce manuel répond en partie à cet obstacle, et la présente édition française poursuit le même objectif pour la filière francophone.

Le tableau 14.1 transforme ces cadres en liste de vérification. Chaque ligne décrit une action que l'on peut observer et noter, et non une attitude générale. L'étudiant qui peut cocher toutes les lignes possède la base exigée à l'examen pratique (TP 14).

Tableau 14.1. Liste des compétences numériques à la fin de la première année, alignée sur les cadres du chapitre 1. Source : adapté de [29,30,34,35].

Domaine (DigComp)	L'étudiant est capable de...	Où il s'est exercé	Domaine DECODE
Information et données	construire une requête PubMed avec MeSH, opérateurs et filtres, et vérifier une référence par son PMID ou son DOI	Chap. 1 ; TP 2	science des données de santé
Information et données	classer une donnée selon sa structure et son échelle, et rédiger un dictionnaire de données	Chap. 2, 3 ; TP 3	science des données de santé
Information et données	coder un diagnostic en CIM-10 et en CIM-11, en indiquant la version utilisée	Chap. 4 ; TP 4	systèmes d'information
Communication et collaboration	partager un document avec le rôle minimal et rédiger une note de téléconsultation	TP 1, TP 9	professionnalisme numérique
Communication et collaboration	reconnaître le standard approprié (HL7 v2, FHIR, DICOM, IPS) dans un scénario	Chap. 7, 8 ; TP 8	systèmes d'information
Création de contenu	rédiger un texte avec des citations Vancouver et la déclaration relative à l'IA	Chap. 5 ; TP 5	professionnalisme numérique
Sécurité	choisir un mot de passe long, activer la validation en deux étapes et reconnaître l'hameçonnage	Chap. 11 ; TP 1, TP 11	professionnalisme numérique
Sécurité	appliquer le RGPD aux données des patients : catégorie particulière, pseudonymisation, notification	Chap. 5	santé du patient et de la population
Résolution de problèmes	mesurer en millimètres sur une image DICOM calibrée et calculer SDNN et RMSSD	Chap. 6, 7 ; TP 6, TP 7	science des données de santé
Résolution de problèmes	calculer Se, Sp, VPP, VPN à différentes prévalences et vérifier une réponse d'un LLM	Chap. 10, 13 ; TP 10, TP 13	science des données de santé

La liste a une limite évidente. Elle vérifie des savoir-faire d'utilisateur, et non des compétences de spécialiste. L'étudiant doit toutefois savoir quand il a besoin d'un spécialiste et quelles questions lui poser.

14.3 Le calendrier des réglementations 2024-2031 qui changeront la pratique

L'informatique médicale de 2026 ne peut pas être enseignée sans calendrier. Bon nombre des règles abordées dans l'ouvrage ont été adoptées, mais s'appliquent de manière échelonnée, et certaines échéances ont déjà été modifiées. La figure 14.2 et le tableau 14.2 rassemblent les dates vérifiées au 27 septembre 2026.

Intelligence artificielle. Le règlement (UE) 2024/1689 (AI Act) est entré en vigueur le 1er août 2024. Les interdictions de l'article 5 et l'obligation de maîtrise de l'IA prévue à l'article 4 s'appliquent depuis le 2 février 2025, et les règles relatives aux modèles d'IA à usage général, à la gouvernance et aux sanctions depuis le 2 août 2025 [22]. L'omnibus numérique sur l'IA, le règlement (UE) 2026/1744 du 8 juillet 2026, est entré en vigueur le 27 juillet 2026 et a décalé les échéances applicables aux systèmes à haut risque. Pour les systèmes autonomes de l'annexe III, parmi lesquels le triage des patients aux urgences, les obligations s'appliquent à partir du 2 décembre 2027. Pour les systèmes intégrés dans des produits relevant de l'annexe I, donc aussi pour les dispositifs médicaux dotés d'IA, elles s'appliquent à partir du 2 août 2028 [197]. Le même règlement a reformulé l'article 4 : les fournisseurs et les déployeurs prennent des mesures pour soutenir la maîtrise de l'IA de leur personnel, sans devoir garantir un niveau déterminé pour chaque personne [197]. Le MDR et l'IVDR restent à l'annexe I, section A [526]. La proposition de la Commission de décembre 2025 visant à les faire passer dans la section B n'était pas adoptée à la date de rédaction [524].

Données de santé. Le règlement (UE) 2025/327 relatif à l'espace européen des données de santé (EHDS) s'applique, en règle générale, à partir du 26 mars 2027. Le résumé du patient, les prescriptions et les dispensations électroniques entrent dans le régime obligatoire le 26 mars 2029, et l'imagerie, les résultats de laboratoire et les rapports de sortie le 26 mars 2031. L'utilisation secondaire, par l'intermédiaire des organismes responsables de l'accès aux données de santé, s'applique elle aussi à partir de 2029 [23,319]. Le Data Act s'applique depuis le 12 septembre 2025, et l'obligation, pour les produits connectés, de permettre l'accès aux données qu'ils génèrent concerne les produits mis sur le marché après le 12 septembre 2026 [527]. En revanche, la proposition "Digital Omnibus", qui modifierait le RGPD, n'était pas adoptée lors de la mise à jour du 1er août 2026 [528]. Le RGPD s'applique donc sous sa forme connue.

Sécurité et identité numérique. La directive NIS2 s'applique depuis le 18 octobre 2024, et la santé y figure parmi les secteurs hautement critiques. Un incident important se déclare en trois temps : alerte précoce dans les 24 heures, notification dans les 72 heures, rapport final dans un délai d'un mois au plus [188]. La Roumanie a transposé la directive par l'ordonnance d'urgence du Gouvernement (OUG) n° 155/2024, publiée le 31 décembre 2024, qui désigne la DNSC comme autorité compétente et comme équipe nationale de réponse aux incidents [187]. Le règlement eIDAS 2 est en vigueur depuis le 20 mai 2024 [434], et la Commission européenne précise que les portefeuilles européens d'identité numérique seront disponibles d'ici la fin de l'année 2026 [437]. La loi n° 214/2024 a remplacé la loi n° 455/2001 relative à la signature électronique [436].

Classifications. La CIM-11 est en vigueur depuis le 1er janvier 2022, et l'aide-mémoire de l'OMS de mai 2026 fait état de 132 pays à différents stades de mise en œuvre [140,142]. La version actuelle du navigateur est la CIM-11 MMS, version 2026-01 [141]. En Roumanie, les diagnostics des patients hospitalisés sont encore codés en ICD-10-AM, 3e édition, pour le groupage RO DRG v.1.1 [139]. Nous n'avons identifié ni acte normatif ni calendrier officiel pour le passage à la CIM-11.

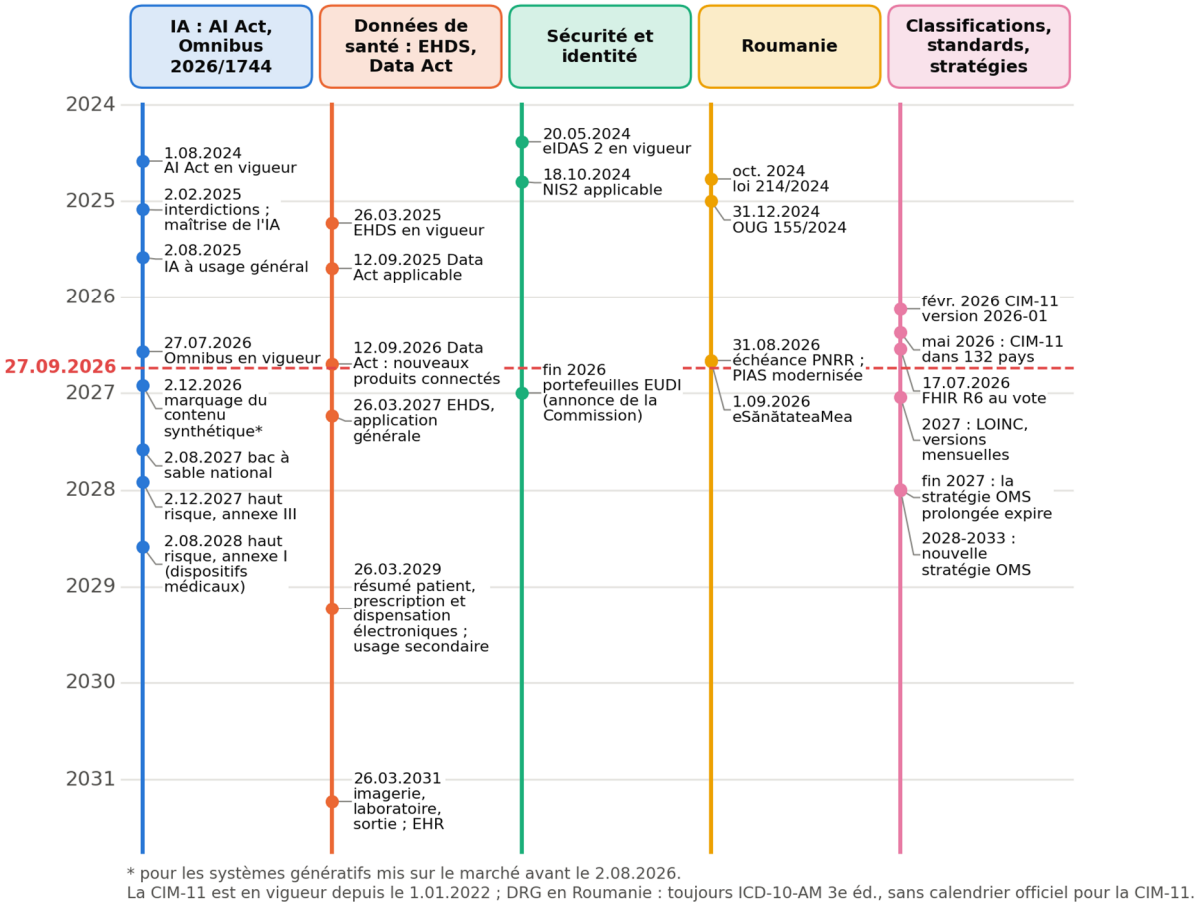


Figure 14.2. Calendrier des réglementations et des standards qui modifient la pratique médicale, 2024-2031, avec la ligne de la date de référence de l'ouvrage (27 septembre 2026). Données issues de [22,23,38,142,187,188,197,307,323,324,434,436,437,527,529].

Tableau 14.2. Calendrier des réglementations 2024-2031 et conséquences pour le médecin. Source : [22,23,187,188,197,434,437].

Date	Acte	Ce qui s'applique	Ce que cela signifie pour le médecin
20.05.2024	eIDAS 2 (2024/1183)	cadre du portefeuille européen d'identité	identification numérique reconnue dans l'UE
18.10.2024	NIS2 (2022/2555)	cybersécurité dans les secteurs critiques, y compris la santé	l'incident se déclare dans les 24 h, 72 h, un mois
31.12.2024	OUG 155/2024	transposition de NIS2 en Roumanie	la DNSC est l'autorité compétente
02.02.2025	AI Act, art. 4 et 5	pratiques interdites, maîtrise de l'IA	l'établissement soutient la formation du personnel
27.07.2026	Règlement 2026/1744	nouveau calendrier pour le haut risque	les échéances de 2026 ne sont plus valables
26.03.2027	EHDS (2025/327)	application générale	institutions, points de contact, organismes d'accès
02.12.2027	AI Act, annexe III	haut risque, systèmes autonomes	par exemple le triage des patients aux urgences
02.08.2028	AI Act, annexe I	haut risque dans les produits réglementés	dispositifs médicaux dotés d'IA (MDR)
26.03.2029	EHDS	résumé, prescription et dispensation électroniques ; utilisation secondaire	échange obligatoire via MyHealth@EU
26.03.2031	EHDS	imagerie, laboratoire, sortie ; systèmes de DSE	l'ensemble du dossier prioritaire devient accessible au patient

La situation en Roumanie en 2026

- Le 1er septembre 2026, la CNAS a lancé le portail eSănătateaMea, avec authentification par ROeID, les prescriptions des 5 dernières années et l'historique de santé des 10 dernières années [324].
- Le projet PNRR de modernisation de la PIAS a été déclaré achevé le 31 août 2026, date limite européenne pour les jalons et les cibles [323,468].
- La Roumanie n'était pas membre de SNOMED International au 27 septembre 2026 [151].
- L'état d'avancement du portefeuille européen d'identité numérique en Roumanie n'a pas pu être confirmé par des sources officielles à la date de rédaction.

14.4 L'éthique de l'utilisation de l'IA par l'étudiant et le médecin

Les chapitres 5, 10 et 13 ont abordé l'IA sous différents angles : la protection des données, la consultation avec un modèle de langage et l'évaluation des modèles. Les principes qui en découlent sont les mêmes pour l'étudiant qui rédige un exposé et pour le médecin qui utilise un outil dans son service.

Confidentialité. En Roumanie, les informations relatives à l'état du patient restent confidentielles même après son décès [171]. Les données concernant la santé constituent une catégorie particulière, dont le traitement est interdit en règle générale et n'est autorisé que dans les exceptions prévues à l'article 9 du RGPD [118]. Le texte saisi dans un chatbot public aboutit sur les serveurs du fournisseur. C'est pourquoi on ne saisit dans les outils publics que des cas fictifs ou des données véritablement anonymisées, et se limiter aux "initiales" n'anonymise rien. La loi n° 190/2018 ajoute une condition pour les décisions automatisées fondées sur des données de santé : un consentement explicite ou des dispositions légales expresses, assortis de mesures de protection [176].

Vérification. Un modèle de langage ne consulte pas une base de faits et ne sait pas quand il se trompe. Sur 636 références, 55 % de celles générées par GPT-3.5 et 18 % de celles générées par GPT-4 n'existaient pas [408]. Lors de la recherche documentaire pour 11 revues systématiques, les taux d'hallucination ont été de 39,6 %, 28,6 % et 91,4 % pour GPT-3.5, GPT-4 et Bard [44]. Même dans un cadre contrôlé de génération de notes cliniques, 1,47 % des phrases contenaient des hallucinations et 3,45 % des omissions, tandis que 44 % des phrases comportant des hallucinations présentaient des erreurs majeures [406]. La règle pratique reste celle du chapitre 1 : chaque affirmation se vérifie dans une recommandation de pratique clinique ou dans une source primaire, et chaque référence se recherche dans PubMed ou CrossRef.

Responsabilité. La performance du modèle seul ne garantit pas celle de l'humain qui l'utilise. Dans l'essai randomisé de Bean et ses collaborateurs, trois modèles ont identifié seuls l'affection pertinente dans 94,9 % des scénarios, mais les participants qui les ont utilisés y sont parvenus dans moins de 34,5 % des cas [403]. À l'inverse, une suggestion erronée peut faire chuter même un expert : lorsque le système proposait une catégorie BI-RADS erronée, la proportion de lectures correctes des radiologues est passée d'environ 80 % à 20-46 %, selon l'expérience [281]. Le code de conduite de l'Académie nationale de médecine des États-Unis exige une responsabilité identifiable et des contrôles portant sur les conséquences indésirables [521]. La décision et la signature restent celles du médecin.

Déclaration. Les recommandations de l'ICMJE mises à jour en janvier 2026 exigent la déclaration de l'utilisation de l'IA : l'aide à la rédaction dans la section des remerciements,

l'utilisation pour les données, l'analyse ou les figures dans la section des méthodes, avec l'outil et sa version. Un chatbot ne peut pas être auteur, et le contenu généré par l'IA ne se cite pas comme source primaire [202]. Le COPE a adopté la même position [203]. Les détecteurs de texte généré ne résolvent pas le problème. Ils classent fréquemment comme générés les textes rédigés par des locuteurs non natifs de l'anglais [205], donc aussi ceux des étudiants roumains qui écrivent en anglais.

Équité. Un algorithme peut être exact en moyenne et inéquitable pour certains sous-groupes. L'hypoxémie occulte, non détectée par l'oxymètre de pouls, est apparue dans 11,7 % des mesures chez les patients noirs, contre 3,6 % chez les patients blancs [192]. Les classificateurs de radiographies thoraciques ont sous-diagnostiqué de façon sélective les femmes, les patients noirs et les patients de faible statut socio-économique [194]. Les recommandations STANDING Together demandent de documenter la composition des jeux de données et d'évaluer la performance par groupe de population [196]. Une revue systématique de l'éthique des modèles de langage en médecine a retrouvé les mêmes préoccupations récurrentes (équité, biais, non-malfaisance, transparence, confidentialité) et la même exigence de supervision humaine [409].

14.4.1 Liste de décision : puis-je utiliser un outil d'IA pour cette tâche ?

Les principes ci-dessus se traduisent en cinq questions à poser avant toute utilisation (figure 14.3). L'ordre compte, car les deux premières questions peuvent arrêter la tâche avant même que l'outil soit ouvert.

1. Est-ce que je saisis des données qui peuvent identifier un patient ? Si oui, je m'arrête. Je travaille sur un cas fictif ou j'utilise uniquement l'outil approuvé par l'établissement pour les données cliniques.
2. Les règles du contexte autorisent-elles l'IA ? La discipline, l'examen, la revue ou l'hôpital peuvent en interdire l'utilisation. À l'examen pratique, l'IA ne s'utilise que si le sujet le demande explicitement.
3. Puis-je vérifier chaque affirmation et chaque référence dans une source indépendante ? Sinon, je ne reprends pas le contenu.
4. La décision ou le texte final restent-ils les miens, et puis-je les justifier sans l'outil ? Sinon, je n'utilise pas l'outil pour cette tâche.
5. Le résultat peut-il désavantager un groupe de patients ? Si oui, je vérifie la performance par sous-groupe et je signale la limite.

Si toutes les réponses permettent de continuer, on utilise l'outil, on vérifie le résultat et on déclare l'utilisation en précisant l'outil, la version et l'objectif.

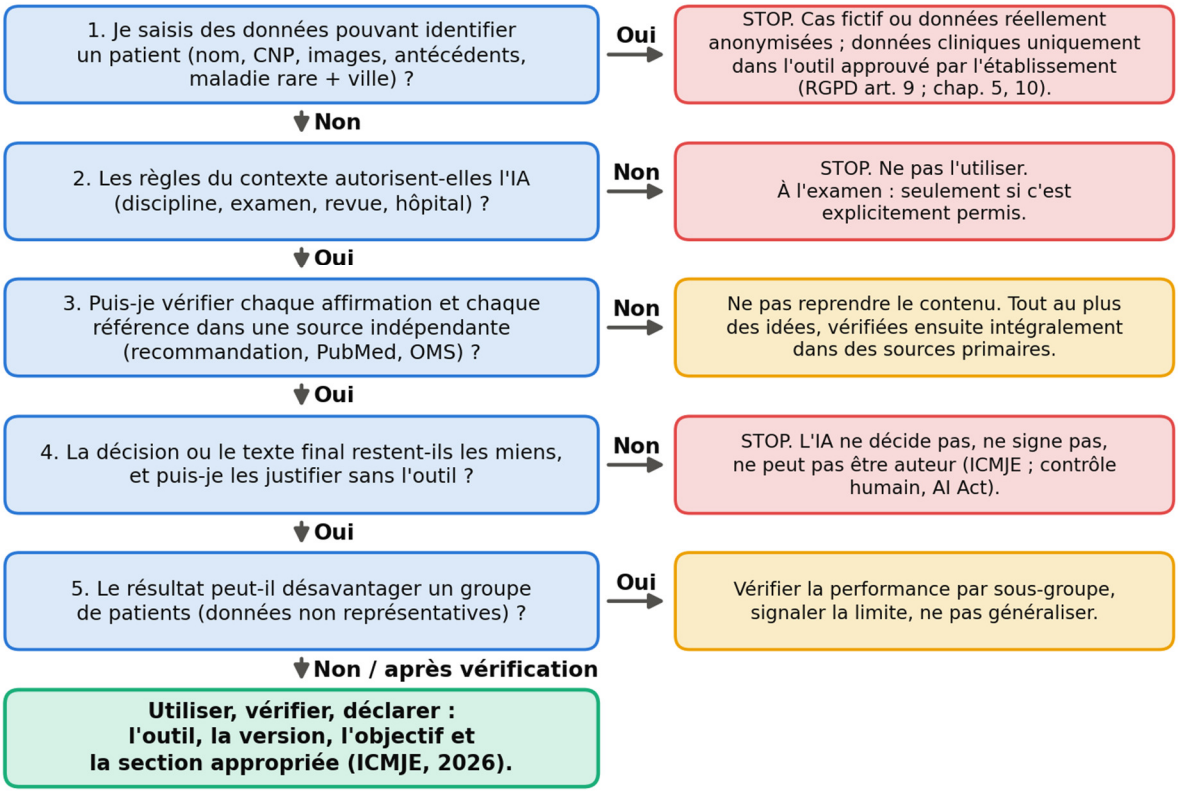


Figure 14.3. La liste de décision "Puis-je utiliser un outil d'IA pour cette tâche ?" sous forme de schéma. Les sorties rouges arrêtent la tâche, les sorties jaunes exigent des mesures supplémentaires. Élaboré d'après [22,118,196,202].

14.5 Orientations après 2026

Quelques orientations peuvent être décrites avec prudence, à partir de documents déjà publiés. La première est la réutilisation des données. À partir de 2029, les chercheurs demanderont l'accès aux données de santé par l'intermédiaire d'organismes nationaux connectés à HealthData@EU, et les données ne seront traitées que dans des environnements sécurisés [23]. La qualité des données collectées aujourd'hui dans les hôpitaux déterminera ce qu'il sera possible d'étudier alors.

La deuxième concerne les modèles. Moor et ses collaborateurs ont proposé le concept d'IA médicale généraliste, capable de combiner images, données du DSE, résultats d'analyses, génomique et texte, et prévoient que de tels modèles mettront à l'épreuve les modes actuels de réglementation et de validation [487]. La validation ne s'achève toutefois jamais. Van Calster et ses collaborateurs soutiennent que les populations, les procédures de mesure et la pratique évoluent, de sorte qu'aucun modèle n'est validé une fois pour toutes [495]. L'exemple du modèle de sepsis désactivé en 2020, après que la pandémie a modifié la relation entre la fièvre et le sepsis bactérien, montre pourquoi la surveillance après la mise en œuvre devient une tâche permanente [496]. Le rapport de la NAM sur l'IA générative recommande précisément une surveillance continue [497].

La troisième concerne le marché. La liste de la FDA des dispositifs médicaux dotés d'IA comptait 1 614 entrées en septembre 2026, dont 76,2 % en radiologie [267]. La part de la radiologie a néanmoins diminué, passant de 85,5 % en 2020-2022 à 77,5 % en 2023-2025, et 96,2 % des dispositifs analysés ont été autorisés par la procédure 510(k), fondée sur l'équivalence [268].

La quatrième concerne les standards. Au 27 septembre 2026, FHIR R6 était encore au stade du vote [307], LOINC annonce le passage à des publications mensuelles en 2027 [529], et la stratégie mondiale de l'OMS pour la santé numérique a été prolongée jusqu'en 2027, une nouvelle stratégie étant demandée pour 2028-2033 [38]. Tout chiffre de cet ouvrage devra donc être vérifié de nouveau dans un an.

Les méthodes de recherche avancées, comme les jumeaux numériques et l'apprentissage fédéré, sont traitées dans la monographie doctorale de l'auteur.

14.6 Questions de révision

Les questions ci-dessous ajoutent, après les cinq questions du chapitre, 20 questions couvrant les chapitres 1-13, au format du concours de l'internat (rezidențiat roumain) : questions à complément simple (QCS), avec une seule bonne réponse, et questions à complément multiple (QCM), avec 2 à 4 bonnes réponses. Chaque réponse peut être retrouvée dans le texte du chapitre indiqué entre parenthèses. Les réponses figurent sur la dernière ligne du chapitre et dans l'annexe A.

Pièges fréquents

- Utiliser les échéances initiales de l'AI Act pour les systèmes à haut risque. Depuis le règlement 2026/1744, les dispositifs médicaux dotés d'IA entrent dans ce régime à partir du 2 août 2028.
- Croire que l'EHDS s'applique intégralement depuis 2025. Le règlement est en vigueur depuis le 26 mars 2025, mais ses obligations s'appliquent de manière échelonnée, en 2027, 2029 et 2031.
- Supposer que la Roumanie a adopté la CIM-11. Au 27 septembre 2026, les DRG reposent encore sur l'ICD-10-AM, 3e édition.
- Considérer comme anonyme un cas saisi dans un chatbot "avec les initiales seulement".
- Traiter une référence proposée par l'IA comme vérifiée parce qu'elle comporte un PMID. Le PMID peut être inventé ou appartenir à un autre article.
- Confondre la performance du modèle testé seul avec celle du médecin ou du patient qui l'utilise.
- Apprendre les chiffres de l'ouvrage comme des vérités immuables. Les standards, les plateformes et les échéances changent d'une année à l'autre.

À retenir

- La discipline comporte cinq couches : données, standards, systèmes, utilisation clinique, gouvernance. Chaque couche dépend de celles qui se trouvent en dessous.
- Les compétences de la première année sont celles d'un utilisateur : recherche documentaire, codage, protection des données, sécurité, mesure, calcul de performance, vérification de l'IA.
- AI Act : interdictions et maîtrise de l'IA depuis le 2 février 2025, haut risque à partir du 2 décembre 2027 (annexe III) et du 2 août 2028 (dispositifs médicaux).
- EHDS : application générale à partir du 26 mars 2027, résumé et prescription électronique à partir de 2029, imagerie et laboratoire à partir de 2031.
- NIS2 et OUG 155/2024 : l'incident important se déclare dans les 24 heures, dans les 72 heures et dans un délai d'un mois.
- Avant d'utiliser l'IA : pas de données identifiables, une autorisation, une vérification, une décision personnelle, une attention à l'équité, puis la déclaration.

Questions d'autoévaluation

1. (QCS) Selon le règlement (UE) 2026/1744, les obligations relatives aux systèmes d'IA à haut risque intégrés dans des produits relevant de l'annexe I, y compris les dispositifs médicaux, s'appliquent à partir du :

a) 2 février 2025
b) 2 août 2026
c) 2 décembre 2027
d) 2 août 2028
e) 26 mars 2031

2. (QCS) L'obligation relative à la maîtrise de l'IA (art. 4 de l'AI Act) s'applique à partir du :

a) 1er août 2024
b) 2 février 2025
c) 2 août 2025
d) 27 juillet 2026
e) 2 décembre 2027

3. (QCS) Au 27 septembre 2026, le groupage DRG des cas hospitalisés en Roumanie repose sur des diagnostics codés en :

a) CIM-11 MMS, version 2026-01
b) ICD-10-AM, 3e édition
c) ICD-10-CM
d) SNOMED CT
e) ICD-9-CM

4. (QCM) Parmi les règles suivantes d'utilisation de l'IA par l'étudiant et le médecin, lesquelles sont étayées dans le chapitre ?

a) on ne saisit pas dans un chatbot public de données pouvant identifier un patient
b) toute référence proposée par un modèle de langage se vérifie dans PubMed ou CrossRef
c) un chatbot peut être coauteur s'il a rédigé une grande partie du texte
d) l'utilisation de l'IA se déclare, avec l'outil, la version et l'objectif
e) la performance d'un outil se vérifie aussi par sous-groupe de patients

5. (QCM) Parmi les compétences suivantes, lesquelles correspondent au domaine DigComp "Sécurité", tel qu'il est appliqué dans le tableau 1.2 et dans le tableau 14.1 ?

a) l'activation de la validation en deux étapes
b) le choix d'un mot de passe long, selon les recommandations du NIST
c) la construction d'une requête PubMed avec MeSH
d) l'application des règles du RGPD aux données des patients
e) le calcul du SDNN à partir d'une série d'intervalles RR

Questions de révision

6. (QCS) (Chap. 1) Le cours d'informatique médicale est devenu obligatoire pour tous les étudiants en médecine de Roumanie en :

a) 1977
b) 1984
c) 1990
d) 1994
e) 2009

7. (QCS) (Chap. 2) Un ECG de 10 secondes, à 12 dérivations, échantillonné à 500 Hz et stocké sur 16 bits, occupe environ :

a) 12 ko
b) 120 ko
c) 1,2 Mo
d) 12 Mo
e) 250 Mo

8. (QCM) (Chap. 3) À propos des clés et de la normalisation, les affirmations correctes sont :

a) la clé primaire identifie de manière unique chaque ligne d'une table
b) une cellule contenant la valeur "I10, E11.9" respecte la première forme normale
c) la clé étrangère doit désigner une ligne existante d'une autre table
d) le nom du patient est une clé primaire appropriée
e) en troisième forme normale, l'année de naissance est déplacée de la table Consultation vers la table Patient

9. (QCS) (Chap. 4) Dans la CIM-11, le séparateur qui relie deux codes souches dans un cluster est :

a) &
b) /
c) †
d) *
e) le point

10. (QCM) (Chap. 4, 12) À propos du codage et des DRG en Roumanie, au 27 septembre 2026, les affirmations correctes sont :

a) les diagnostics des patients hospitalisés sont codés en ICD-10-AM, 3e édition
b) la classification DRG en vigueur est la RO DRG v.1.1, de mars 2024
c) l'INMSS collecte auprès des hôpitaux l'ensemble minimal de données au niveau du patient
d) la CIM-11 a officiellement remplacé l'ICD-10-AM depuis 2022
e) les actes sont codés en CIM-11

11. (QCM) (Chap. 5) Selon la directive NIS2 et l'OUG n° 155/2024, un incident important de cybersécurité se déclare par :

a) une alerte précoce dans les 24 heures
b) une notification dans les 72 heures
c) un rapport final dans un délai d'un mois au plus
d) une déclaration uniquement si des données personnelles sont touchées
e) une déclaration exclusivement auprès de l'ANSPDCP

12. (QCS) (Chap. 6) Un signal comporte des composantes utiles jusqu'à 150 Hz. D'après le théorème d'échantillonnage, la fréquence d'échantillonnage doit être :

a) d'au moins 75 Hz
b) d'exactement 150 Hz
c) supérieure à 300 Hz
d) de 256 Hz, quel que soit le signal
e) inférieure à 50 Hz, pour éviter l'interférence du réseau électrique

13. (QCS) (Chap. 7) La fenêtre pulmonaire, de niveau L = -600 UH et de largeur W = 1 500 UH, affiche sur toute l'échelle de gris l'intervalle :

a) de -1 350 à +150 UH
b) de -600 à +900 UH
c) de -1 500 à 0 UH
d) de -160 à +240 UH
e) de -2 100 à +900 UH

14. (QCM) (Chap. 7) À propos du standard DICOM, les affirmations correctes sont :

a) une étude correspond approximativement à une procédure d'imagerie demandée
b) chaque instance possède un identifiant unique (UID)
c) l'export au format JPEG conserve la calibration spatiale (Pixel Spacing)
d) le service C-FIND recherche des objets selon des critères
e) la fenêtre d'affichage modifie les valeurs UH du fichier

15. (QCS) (Chap. 8) La version de HL7 FHIR sur laquelle repose la majorité des implémentations actuelles, et qu'utilisent le guide IPS et le serveur public de test du TP 8, est :

a) R2
b) R3
c) R4
d) R6
e) R7

16. (QCM) (Chap. 8, 10) À propos du règlement (UE) 2025/327 relatif à l'EHDS, les affirmations correctes sont :

a) il s'applique, en règle générale, à partir du 26 mars 2027
b) les obligations relatives au résumé du patient, à la prescription électronique et à la dispensation électronique s'appliquent à partir du 26 mars 2029
c) l'imagerie, les résultats de laboratoire et les rapports de sortie s'y ajoutent à partir du 26 mars 2031
d) il doit être transposé en Roumanie par une loi nationale
e) les systèmes de DSE doivent comporter un composant européen de journalisation

17. (QCS) (Chap. 9) Les normes méthodologiques de la télémédecine en Roumanie ont été approuvées par :

a) l'arrêté n° 1133/2021
b) la décision du Gouvernement n° 1133/2022
c) la loi n° 45/2019
d) l'OUG n° 155/2024
e) la loi n° 214/2024

18. (QCM) (Chap. 9) À propos des données probantes sur la télémédecine, les affirmations correctes sont :

a) la télésurveillance dans l'insuffisance cardiaque a été associée à une réduction de la mortalité toutes causes (OR 0,84)
b) la consultation à distance utilisée seule n'a pas apporté de bénéfice dans l'insuffisance cardiaque, dans une méta-analyse limitée aux maladies cardiovasculaires
c) la télésurveillance tensionnelle a réduit la pression systolique de plusieurs dizaines de mmHg
d) le télé-AVC (*telestroke*) a donné des résultats similaires à ceux de la prise en charge conventionnelle, dans une méta-analyse portant sur 12 540 patients
e) la télépsychiatrie a été supérieure à la consultation en face à face pour tous les diagnostics

19. (QCS) (Chap. 10) Selon la règle 11 du MDR, un logiciel qui fournit des informations utilisées pour prendre des décisions diagnostiques ou thérapeutiques, sans présenter les risques qui le feraient passer dans une classe supérieure, relève de la classe :

a) I
b) IIa
c) IIb
d) III
e) ce n'est pas un dispositif médical

20. (QCM) (Chap. 10, 13) À propos des grands modèles de langage, les données présentées dans l'ouvrage montrent que :

a) les participants qui ont utilisé un LLM ont identifié l'affection pertinente dans moins de 34,5 % des scénarios
b) 18 % des références générées par GPT-4 n'existaient pas, dans une étude portant sur 636 références
c) dans les notes cliniques générées automatiquement, le taux d'hallucinations a été de 1,47 %
d) les chatbots populaires ont fréquemment refusé de répondre aux questions portant sur des sujets propices à la désinformation
e) le score obtenu aux examens d'autorisation d'exercice garantit la sécurité en consultation

21. (QCS) (Chap. 11) Selon eIDAS, l'effet juridique équivalent à celui d'une signature manuscrite est reconnu à :

a) la signature électronique simple
b) la signature électronique avancée
c) la signature électronique qualifiée
d) l'image numérisée de la signature
e) le nom tapé à la fin d'un e-mail

22. (QCM) (Chap. 11) À propos de la cryptographie et de la signature numérique, les affirmations correctes sont :

a) AES est un algorithme de chiffrement symétrique
b) la fonction de hachage n'utilise pas de clé et ne permet pas de reconstituer le document
c) lors de la signature, on signe l'empreinte du document avec la clé privée
d) la vérification de la signature exige la clé privée du signataire
e) TLS protège les données stockées sur le serveur

23. (QCS) (Chap. 12) Un hôpital dispose de 10 applications qui communiquent par l'intermédiaire d'un moteur d'interfaces, dans une architecture en étoile. Le nombre de liaisons nécessaires est de :

a) 10
b) 20
c) 45
d) 90
e) 100

24. (QCS) (Chap. 12) L'indice de case-mix d'un hôpital est :

a) la moyenne des poids relatifs des cas sortis au cours d'une période
b) le nombre de journées d'hospitalisation divisé par le nombre de sorties
c) le nombre de lits pour 100 000 habitants
d) la proportion de réhospitalisations à 30 jours
e) le nombre de groupes DRG utilisés par l'hôpital

25. (QCM) (Chap. 13) Un outil d'IA a une sensibilité de 0,80 et une spécificité de 0,95. Les affirmations correctes sont :

a) la valeur prédictive positive diminue lorsque la prévalence dans la population testée diminue

b) la sensibilité et la spécificité ne dépendent pas, en principe, de la prévalence

c) une aire sous la courbe ROC de 0,90 garantit une valeur prédictive positive d'au moins 0,90

d) l'exactitude peut être trompeuse lorsque la maladie est rare

e) le score F1 tient compte des vrais négatifs

PARTIE II

Partie II. Travaux pratiques

Travaux pratiques 1. L'ordinateur, le système d'exploitation et l'espace de travail collaboratif (Google Workspace)

Objectif

À la fin de la séance, l'étudiant organise ses fichiers de cours dans une arborescence de dossiers aux noms cohérents, partage un document avec des droits minimaux, utilise l'historique des versions, active la validation en deux étapes sur son propre compte et explique pourquoi les données des patients n'ont pas leur place dans un compte cloud personnel. Tous les exemples de patients de cette séance sont fictifs.

Outils

- L'ordinateur de la salle de TP, sous Windows 11, ou l'ordinateur personnel (Windows, macOS, ChromeOS ou une distribution Linux).
- Un navigateur à jour (Chrome, Edge ou Firefox, version en vigueur au 27 septembre 2026).
- Un compte Google personnel ou le compte institutionnel, si la faculté en fournit un. Google Drive, Docs et Sheets sont des applications web, sans numéro de version affiché. Les pages d'aide citées ont été consultées le 27 septembre 2026 [530,531].
- En option, un gestionnaire de mots de passe de confiance [532].

L1.1 L'ordinateur et les fichiers : notions minimales

Un ordinateur comporte deux niveaux. Les composants physiques (*hardware*) sont le processeur, qui exécute les instructions, la mémoire vive (RAM), qui conserve temporairement les données des programmes ouverts, l'unité de stockage (SSD ou HDD), où les données restent après l'arrêt, et les périphériques d'entrée et de sortie. Les programmes (*software*) se répartissent entre le système d'exploitation et les applications. Le système d'exploitation gère les ressources physiques, les fichiers et les utilisateurs. Les applications, comme le traitement de texte ou le tableur, fonctionnent au-dessus de lui. C'est aussi le système d'exploitation qui sépare les comptes des utilisateurs, de sorte que chacun ne voit que ses propres fichiers, et qui reçoit des mises à jour corrigeant, entre autres, des vulnérabilités de sécurité. Un ordinateur partagé, comme celui de la salle de TP, exige donc deux habitudes simples : se déconnecter de son compte en partant et éviter d'enregistrer ses mots de passe dans le navigateur d'un ordinateur utilisé par plusieurs personnes.

Les fichiers sont organisés dans un système hiérarchique de dossiers. Chaque fichier a un nom et, en règle générale, une extension qui indique le format : `.docx`, `.pdf`, `.csv`. Le format détermine quel programme peut ouvrir le fichier et avec quelle facilité il pourra être lu dans dix ans. En médecine, le choix du format a des conséquences pratiques. Une image radiologique exportée en JPEG perd les métadonnées que conserve le format DICOM, et une table d'étude enregistrée

uniquement dans le format d'un logiciel commercial peut devenir difficile à ouvrir après un changement de licence. Le standard DICOM a été créé pour les images et les tracés, et HL7 FHIR décrit les données cliniques sous forme de ressources XML ou JSON [15]. Les signaux physiologiques multicanaux, comme l'EEG ou la polysomnographie, s'échangent souvent au format EDF ou EDF+ [221]. Le tableau L1.1 résume les formats que l'étudiant rencontrera dans cet ouvrage.

Tableau L1.1. Formats de fichiers rencontrés dans l'activité médicale et dans les travaux pratiques. Source : [15,221].

Format	Extension	Utilisation typique	Remarque
Document texte	.docx, .odt	lettres, rapports, exposés	modifiable, s'envoie en PDF
Document final	.pdf	résultats, documents signés	mise en page fixe sur tout appareil
Tableau	.xlsx, .ods	données d'étude, calculs	conserve formules et mise en forme
Texte délimité	.csv	échange de données entre programmes	texte brut, sans formules
Image générale	.png, .jpg	figures, photographies	JPEG compresse avec pertes
Imagerie médicale	.dcm (DICOM)	radiologie, TDM, IRM	contient des métadonnées sur le patient
Données cliniques	.json, .xml (HL7 FHIR)	échange entre systèmes informatiques	structuré en ressources
Signaux	.edf (EDF, EDF+)	EEG, polysomnographie, ECG	multicanal, format ouvert

Le nom d'un fichier doit pouvoir être compris sans ouvrir le fichier. La règle proposée dans cette séance est `AAAA-MM-JJ_activite_auteur_vNN.extension`, par exemple `2026-10-06_TP01_Popescu_v02.docx`. La date écrite dans l'ordre année-mois-jour permet aux fichiers de se ranger chronologiquement lorsqu'ils sont triés par nom. Les espaces et les signes diacritiques dans les noms peuvent poser problème lors du transfert entre systèmes, c'est pourquoi on les évite. La version s'écrit à la fin, sur deux chiffres. La figure L1.1 montre une arborescence de dossiers pour le premier semestre.

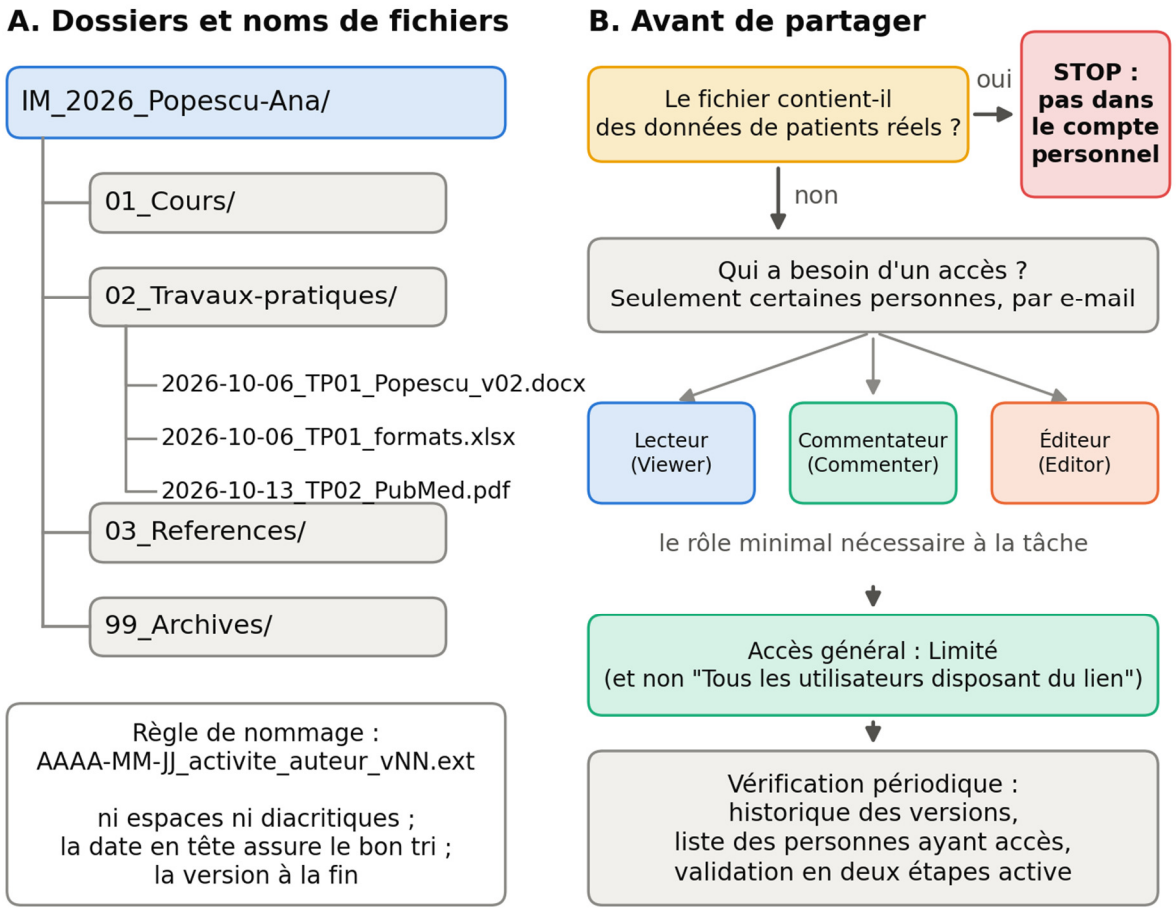


Figure L1.1. Organisation des fichiers de cours (A) et décision de partage (B). Les rôles et les options d'accès sont ceux de Google Drive [530].

L1.2 Partage, versions et sécurité du compte

Google Drive propose quatre rôles. Le lecteur (*Viewer*) peut voir et télécharger le fichier. Le commentateur (*Commenter*) peut en outre ajouter des commentaires. L'éditeur (*Editor*) peut modifier et partager le fichier, mais ne peut pas en changer le propriétaire. Le propriétaire (*Owner*) dispose de tous les droits [530]. Indépendamment des personnes ajoutées nominativement, il existe l'accès général. Avec l'option "Limité" (*Restricted*), seules les personnes ajoutées peuvent ouvrir le fichier. Avec l'option "Tous les utilisateurs disposant du lien" (*Anyone with the link*), le fichier peut être utilisé par quiconque reçoit le lien, sans authentification [530]. Le principe de travail est celui du rôle minimal nécessaire : qui ne fait que lire reçoit le rôle de lecteur.

L'historique des versions indique qui a modifié le document et quand. Pour consulter les versions antérieures, il faut disposer du droit de modification. Une version peut recevoir un nom (au plus 40 versions nommées dans un document et 15 dans une feuille de calcul) et peut être restaurée [531]. La même page avertit que certaines révisions peuvent être fusionnées automatiquement, si bien que l'historique ne remplace pas les copies de sauvegarde [531].

Le compte qui héberge les fichiers doit être protégé. Google recommande des mots de passe d'au moins 12 caractères, différents pour chaque compte important, sans données personnelles faciles à deviner, et suggère d'utiliser un gestionnaire de mots de passe de confiance [532]. Le guide américain NIST SP 800-63B-4, publié en 2025, va plus loin. Les mots de passe utilisés

comme facteur d'authentification unique doivent compter au moins 15 caractères, et ceux qui sont utilisés avec un second facteur, au moins 8. Ce guide interdit les règles de composition imposées (du type "au moins un chiffre et un symbole") et le changement périodique forcé du mot de passe, mais elle exige de le changer lorsqu'il existe des indices de compromission. Les vérificateurs doivent autoriser les gestionnaires de mots de passe et la saisie automatique, et le collage du mot de passe depuis le presse-papiers est recommandé [445]. Une longue phrase de passe, facile à mémoriser et utilisée à un seul endroit, est donc un bon choix.

La validation en deux étapes ajoute un second facteur, utile si le mot de passe a été volé. Google propose plusieurs options : des notifications sur le téléphone, des clés d'accès (*passkeys*) et des clés de sécurité, qui protègent aussi contre les escroqueries par hameçonnage (*phishing*), des applications d'authentification à codes uniques, des codes par SMS ou par appel et des codes de secours à huit chiffres, à conserver en lieu sûr [533].

La sécurité n'est pas une question théorique pour les hôpitaux. Entre 2016 et 2021, 374 attaques par rançongiciel (*ransomware*) contre des prestataires de soins ont été signalées aux États-Unis, exposant les données de santé de près de 42 millions de patients. Le nombre annuel d'attaques a plus que doublé, passant de 43 à 91, et 44,4 % des attaques ont perturbé la prise en charge des patients [183]. Au cours de la même période, les systèmes touchés ont été de moins en moins souvent restaurés à partir de copies de sauvegarde [183]. Une revue systématique a analysé les comportements humains qui ouvrent des brèches de sécurité dans les établissements de santé, notamment la vulnérabilité du personnel aux attaques d'ingénierie sociale [442].

Les copies de sauvegarde (*backup*) se font selon une règle pratique facile à retenir, appelée 3-2-1 : trois copies des données, sur deux types de supports différents, dont une conservée dans un autre lieu. L'agence américaine de cybersécurité recommande de conserver au moins une copie des données critiques hors ligne et chiffrée, et de tester régulièrement la restauration, car de nombreuses variantes de rançongiciels recherchent et effacent ou chiffrent les copies accessibles depuis le réseau [534]. Pour l'étudiant, la solution minimale est la suivante : les fichiers dans Drive, une copie sur l'ordinateur et une copie sur un disque externe gardé déconnecté. La synchronisation automatique ne tient pas lieu de copie de sauvegarde. Si un fichier est supprimé par erreur ou chiffré par un programme malveillant sur l'ordinateur, la modification peut se propager dans le cloud, et la copie synchronisée devient aussi inutilisable que l'original.

La dernière règle est la plus importante pour un futur médecin. Le règlement général sur la protection des données (RGPD) définit les données concernant la santé comme des données à caractère personnel relatives à la santé physique ou mentale d'une personne, y compris la prestation de services de soins de santé. Elles font partie des catégories particulières, dont le traitement est interdit en dehors des exceptions prévues à l'article 9 [118]. Le responsable du traitement, par exemple l'hôpital, doit mettre en œuvre des mesures techniques et organisationnelles appropriées (article 32) et ne peut faire appel à un prestataire de services que sur la base d'un contrat qui oblige celui-ci à protéger les données (article 28) [118]. Le compte personnel d'un étudiant ou d'un médecin n'entre pas dans ce cadre. C'est pourquoi les photographies de documents médicaux, les listes de patients ou les images du moniteur de la chambre ne sont jamais téléversées dans un cloud personnel, même si le dossier est "privé". Le chapitre 5 détaille ces règles.

Déroulement

1. **Les composants de l'ordinateur.** Sur l'ordinateur de la salle de TP, on ouvre Paramètres > Système > Informations système (Windows 11) ou "À propos de ce Mac" (macOS). On note dans le document de travail le type de processeur, la mémoire RAM, la capacité de stockage et la version du système d'exploitation.
2. **Les extensions des fichiers.** Dans l'Explorateur de fichiers, on active l'affichage des extensions (menu Affichage > Afficher, option relative aux extensions de noms de fichiers). On repère trois fichiers portant des extensions différentes et le programme associé à chacun.
3. **L'arborescence de dossiers.** Dans Google Drive, on crée le dossier `IM_2026_Nom-Prenom` et, à l'intérieur, les sous-dossiers `01_Cours`, `02_Travaux-pratiques`, `03_References` et `99_Archives`, comme dans la figure L1.1.
4. **Le premier document.** Dans `02_Travaux-pratiques`, on crée un document Google Docs nommé selon la règle de la section L1.1, par exemple `2026-10-06_TP01_Popescu_v01`. On y inscrit les réponses des étapes 1 et 2.
5. **La feuille de calcul.** Dans le même dossier, on crée une feuille Google Sheets contenant le tableau L1.1, complété par deux autres formats trouvés par l'étudiant (par exemple, le format d'une application de téléphone pour l'ECG). Pour chacun, on note la source de l'information.
6. **Le partage avec un rôle minimal.** Le document est partagé avec un collègue en tant que Commentateur et avec l'enseignant en tant que Lecteur, au moyen du bouton Partager (*Share*). L'accès général reste Limité. L'étudiant vérifie ensuite, dans une fenêtre de navigation privée et sans être connecté, que le lien n'ouvre pas le document.
7. **Commentaire et réponse.** Le collègue ajoute un commentaire à un paragraphe, puis l'auteur répond et marque le commentaire comme résolu.
8. **L'historique des versions.** L'auteur effectue trois modifications successives. Il ouvre l'historique des versions (le bouton en haut à droite qui indique la dernière modification), nomme la version actuelle "après commentaires", puis restaure une version antérieure et revient à la dernière [531]. On note les informations affichées pour chaque version.
9. **La validation en deux étapes.** Sur son propre compte, dans la section sécurité du compte Google, l'étudiant vérifie si la validation en deux étapes est active et, si ce n'est pas le cas, l'active, en choisissant une méthode autre que le SMS lorsque c'est possible [533]. Les codes de secours sont conservés en dehors de l'ordinateur. Aucun mot de passe ni aucun code n'est écrit dans le document de travail.
10. **Le mot de passe.** L'étudiant compare les règles de mot de passe d'un service qu'il utilise aux recommandations du NIST (longueur, pas de règles de composition, pas de changement périodique forcé) et note deux différences [445].
11. **Le plan de sauvegarde.** On décrit, en cinq lignes, le plan personnel 3-2-1 pour le dossier de cours : où se trouvent les trois copies, sur quels supports et à quelle fréquence elles sont mises à jour. On télécharge une copie du document au format PDF et une autre au format DOCX dans le dossier local d'archives.
12. **Scénario RGPD (fictif).** Un interne veut photographier la feuille d'observation du patient fictif "I.P., 67 ans" pour en discuter le soir avec un collègue, via Drive. L'étudiant écrit trois raisons pour lesquelles cela n'est pas permis et propose une solution correcte (par exemple, la discussion à l'hôpital, sur le système de l'établissement, ou un cas entièrement anonymisé, sans détails identifiants).

Travail à rendre

On remet le lien vers le dossier `IM_2026_Nom-Prenom`, partagé avec l'enseignant en tant que Lecteur, avec un accès général Limité. Le dossier contient le document de travail et la feuille de calcul. On ne remet pas de captures d'écran comportant des données du compte personnel. Barème indicatif, 10 points :

- arborescence de dossiers et noms de fichiers conformes à la règle (2 points).
- réponses aux étapes 1 et 2 et tableau des formats complété, avec sources (2 points).
- partage correct : rôles minimaux, accès Limité, vérification dans la fenêtre de navigation privée (2 points).
- historique des versions : version nommée, restauration décrite (1 point).
- comparaison des règles de mot de passe avec celles du NIST et confirmation de la validation en deux étapes, sans codes (1 point).
- plan de sauvegarde et scénario RGPD argumenté (2 points).

À l'examen pratique, on peut demander de refaire les étapes 3, 6 et 8 sur un nouveau document, en 10 minutes.

Variante hors ligne

Si Internet ou les comptes ne sont pas disponibles, la séance se déroule en local. L'arborescence de dossiers est créée dans l'Explorateur de fichiers, sur le disque de l'ordinateur, et les documents sont rédigés dans LibreOffice Writer et Calc ou dans Microsoft Word et Excel. Les versions sont conservées manuellement, par l'enregistrement successif des fichiers `_v01`, `_v02`, `_v03`, et à l'étape 8 l'étudiant compare deux versions à l'aide de la fonction de comparaison de documents du traitement de texte. Le partage est simulé sur papier : pour trois personnes (collègue, enseignant, patient fictif), on remplit un tableau indiquant le rôle accordé et sa justification. La copie de sauvegarde se fait sur une clé USB, et les étapes 9 et 10 sont discutées à partir de la section L1.2. Le dossier est remis sur la clé ou téléversé plus tard, lorsque le réseau est rétabli.

Travaux pratiques 2. L'information médicale sur le web : recherche et évaluation de la qualité

Objectif

L'étudiant transforme une question clinique en une recherche PubMed structurée (PICO, MeSH, opérateurs, filtres), enregistre et exporte les résultats, puis évalue de manière critique trois types de sources : un article de Wikipédia, une page web destinée aux patients et la réponse d'un chatbot (agent conversationnel), comparée à un guide de pratique. La question clinique et le patient des exemples sont fictifs.

Outils

- PubMed, le moteur de recherche gratuit de la Bibliothèque nationale de médecine des États-Unis, et son guide d'utilisation, consultés le 27 septembre 2026 [39,41].
- MeSH Browser, pour vérifier les descripteurs et les synonymes [535].
- Wikipédia, en roumain et en anglais.
- Un chatbot généraliste disponible gratuitement à la date de la séance. On note son nom et la version affichée par l'application.
- Google Sheets ou un autre tableur, pour la liste de vérification du tableau L2.1.

L2.1 De la question clinique à la requête

Une bonne recherche commence par une question bien formulée. Le format PICO décompose la question en population (*Population*), intervention (*Intervention*), comparateur (*Comparator*) et critère de jugement (*Outcome*). Il est utile comme point de départ, mais il ne garantit pas à lui seul une recherche efficace. Une analyse de 2021 montre que les preuves de son efficacité pédagogique sont limitées et recommande de ne pas l'utiliser comme outil exclusif, mais conjointement avec l'évaluation critique de l'ensemble du processus [536].

L'exemple utilisé dans la séance est fictif : "Chez les adultes atteints d'hypertension artérielle, la surveillance de la pression artérielle à distance, par rapport aux soins habituels, fait-elle baisser les valeurs tensionnelles ?" À chaque élément PICO correspond un concept de recherche. Pour chaque concept, on cherche dans MeSH Browser le descripteur approprié : Hypertension, Telemedicine, Blood Pressure [535]. MeSH a une structure hiérarchique, et la recherche par un descripteur inclut automatiquement les termes subordonnés, par explosion. Les synonymes (*entry terms*) montrent sous quels autres noms le concept apparaît dans la littérature [13].

Les articles très récents peuvent ne pas encore être indexés avec MeSH. C'est pourquoi chaque concept est aussi exprimé par des mots du titre et du résumé, reliés au descripteur par OR. Les concepts sont ensuite reliés entre eux par AND. Les opérateurs s'écrivent en majuscules, les parenthèses fixent l'ordre des opérations, l'astérisque tronque le terme (après au moins quatre

lettres), et les guillemets demandent une expression exacte [41]. Pour l'exemple ci-dessus, la requête devient :

```
("Hypertension"[Mesh] OR hypertens*[tiab]) AND ("Telemedicine"[Mesh] OR telemonitor*[tiab] OR "home blood pressure"[tiab]) AND ("Blood Pressure"[Mesh] OR "blood pressure"[tiab])
```

Le 27 septembre 2026, le premier concept seul a renvoyé 669 982 notices, le deuxième 63 552, le troisième 546 995, et la combinaison des trois 3 010 [39]. Le filtre des revues systématiques et des méta-analyses a réduit le résultat à 131 notices, et la limitation supplémentaire aux années 2021-2026, à 73. Les essais contrôlés randomisés de 2021-2026 étaient au nombre de 123 [39]. Ces nombres évolueront avec le temps, c'est pourquoi chaque étudiant note ses propres valeurs et la date de sa recherche.

Pourquoi une simple recherche de type Google ne suffit-elle pas ? Les moteurs généralistes pratiquent la recherche partielle : ils acceptent quelques mots et classent les documents selon leur degré de correspondance avec ces mots. Les bases bibliographiques permettent en outre la recherche exacte, avec des ensembles combinés par des opérateurs booléens et des attributs comme le type de publication [13]. Même lorsqu'il reçoit une phrase ordinaire, PubMed essaie de faire correspondre successivement les mots à des descripteurs MeSH, à des titres de revues, à des expressions fréquentes et à des noms d'auteurs, et il recherche en texte libre ce qui reste sans correspondance [13]. La traduction exacte de la requête s'affiche sur la page Advanced, dans la section d'historique et de détails de la recherche. Elle mérite d'être vérifiée à chaque fois, car elle montre ce que le système a réellement compris. La fonction Clinical Queries applique, en plus des termes de l'utilisateur, des stratégies validées pour les questions de thérapeutique, de diagnostic, d'effets nocifs et de pronostic [13]. Dans la version actuelle de PubMed, les revues systématiques ne font plus partie de Clinical Queries et se recherchent avec le filtre du type d'article [41].

Que faire des résultats ? L'historique des recherches de la page Advanced conserve les 100 dernières requêtes et expire après 8 heures d'inactivité. L'enregistrement permanent de la recherche et les alertes par courriel exigent un compte My NCBI. Les citations s'exportent avec Send to > Citation manager, dans un fichier `.nbib` (au plus 10 000 citations), compatible avec de nombreux logiciels de gestion des références [41]. Le tri par défaut, Best Match, utilise un algorithme de pertinence, et le tri par date reste disponible [41].

Les résultats n'ont pas tous le même poids. La hiérarchie des preuves place les revues systématiques et les méta-analyses au-dessus des essais randomisés, et ceux-ci au-dessus des études observationnelles, des séries de cas et de l'opinion d'experts. La hiérarchie oriente la lecture, mais elle ne dispense pas de l'évaluation critique de chaque article [42]. La figure L2.1 résume l'ensemble du flux.

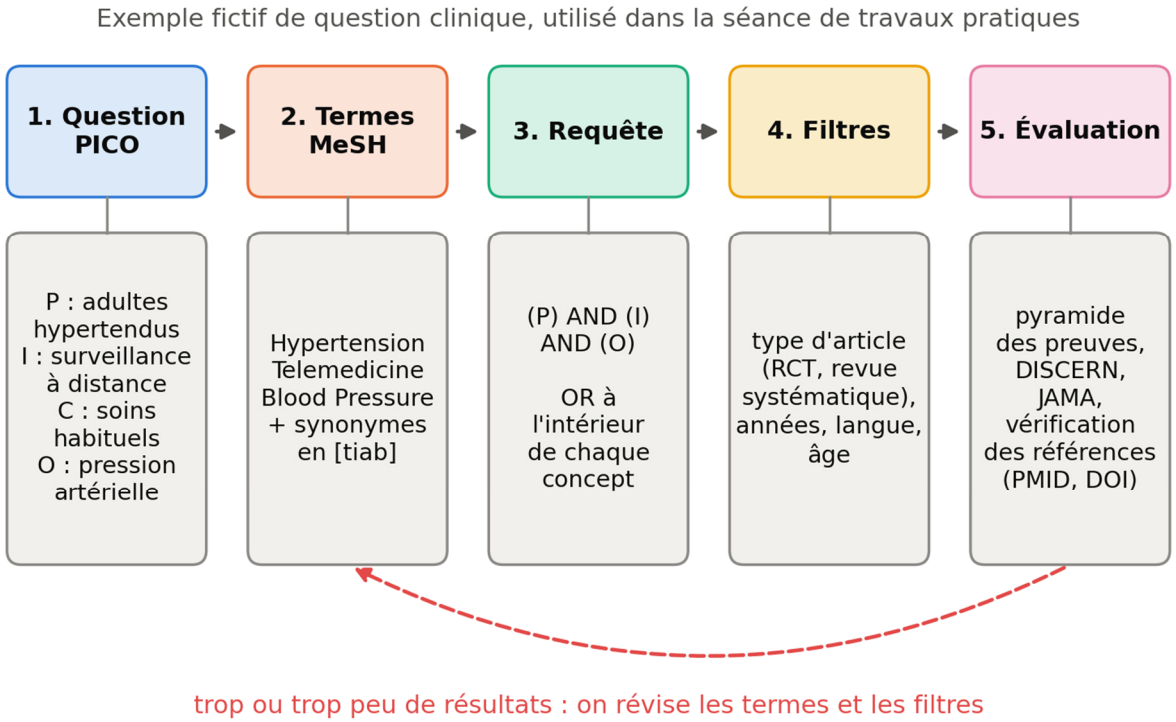


Figure L2.1. Flux de travail, de la question clinique à l'évaluation des sources. Exemple fictif. Syntaxe d'après [41].

L2.2 Wikipédia : un point de départ, pas une source à citer

Wikipédia est l'un des sites les plus visités au monde. La version anglophone contient plus de 45 000 articles médicaux, consultés près de 1,6 milliard de fois par an [537]. Leur qualité est variable. Une étude portant sur les 1 000 articles de santé les plus lus de la Wikipédia anglophone a trouvé une qualité comparable entre les articles sur les affections propres aux femmes et ceux sur les affections propres aux hommes [537]. Les différences entre langues peuvent toutefois être importantes. L'article sur l'accident vasculaire cérébral existait dans 23 des 30 langues les plus parlées, son contenu variait beaucoup d'une langue à l'autre, et trois versions présentaient les corticoïdes comme un traitement de l'accident vasculaire. Le score de qualité était positivement corrélé au nombre de références de l'article [538].

Conclusion pratique : Wikipédia peut servir à l'orientation, à la terminologie et à la liste de références placée à la fin de l'article. Les références doivent être ouvertes et vérifiées. Dans les travaux universitaires, on cite la source primaire ou le guide, et non l'encyclopédie.

L2.3 Outils d'évaluation de l'information destinée aux patients

Deux outils reviennent souvent dans les études sur la qualité de l'information médicale en ligne. DISCERN évalue la qualité de l'information de santé, avec un score total qui peut atteindre 80 points. Les critères du JAMA vérifient quatre éléments : l'auteur, l'attribution des sources, la déclaration des conflits d'intérêts et la date de mise à jour [539,540].

Les résultats de leur application sont en général modestes. Dans une étude portant sur 65 pages web en langue arabe consacrées aux implants dentaires, seules 3 pages (4,5 %) ont obtenu un score DISCERN élevé, et seules 5 (7,6 %) remplissaient tous les critères du JAMA. Les lacunes les plus fréquentes concernaient l'attribution des sources et le nom de l'auteur [539]. Le même type d'évaluation peut s'appliquer aux réponses d'un chatbot. Pour 24 questions formulées à

partir du guide européen d'immunothérapie allergénique, les réponses de ChatGPT-4 ont été de qualité "acceptable" selon DISCERN, mais ont obtenu des scores JAMA de 0-1, car l'auteur, les sources, les déclarations et la date manquaient. Les textes exigeaient un niveau de lecture universitaire [540].

L2.4 La réponse d'un chatbot comparée au guide

Les chatbots répondent avec aisance, mais leurs réponses doivent être vérifiées par rapport à une source de référence. Une étude de 2025 a transformé 28 recommandations d'un guide international sur l'angiœdème héréditaire en questions adressées à quatre modèles (ChatGPT, Gemini, Perplexity, Copilot). Deux cliniciens ont noté l'exactitude, la pertinence, la clarté et la fiabilité des références. ChatGPT et Gemini ont obtenu des scores médians de 5 sur 5 pour l'exactitude et la pertinence, contre 3 pour les deux autres modèles. Lorsqu'on leur a demandé explicitement d'indiquer leurs sources, la cohérence des réponses s'est améliorée [541].

Pour qu'une telle comparaison puisse être reproduite, la méthode de travail doit être documentée. Les lignes directrices de publication CHART, publiées en 2025 pour les études qui évaluent les conseils de santé fournis par des chatbots, comportent 12 items et 39 sous-items. On y trouve l'identification du modèle, la formulation des instructions (*prompt*), la stratégie d'interrogation et le mode d'évaluation de la performance [542]. La séance reprend, de façon simplifiée, la même logique : on note le modèle, la date, le texte exact de la question et la réponse complète.

Tableau L2.1. Liste de vérification pour une source d'information médicale trouvée sur le web ou produite par un chatbot. Source : adapté de [42,539,540,542].

Critère	Question de vérification	Oui / Non / En partie
Auteur	L'auteur est-il nommé et a-t-il une qualification pertinente ?	
Sources	Des sources vérifiables sont-elles citées (PMID, DOI, guide) ?	
Références réelles	Chaque référence existe-t-elle et dit-elle ce qu'on lui attribue ?	
Actualité	La date est-elle affichée et a-t-elle au plus 5 ans ?	
Conflits d'intérêts	Le financement et les intérêts commerciaux sont-ils déclarés ?	
Niveau de preuve	S'appuie-t-elle sur des revues systématiques, des ECR ou des guides ?	
Concordance avec le guide	La recommandation concorde-t-elle avec le guide en vigueur ?	
Équilibre	Les risques et les solutions de rechange sont-ils aussi présentés ?	
Documentation (chatbot)	Le modèle, la date et le texte de la question sont-ils notés ?	

Déroulement

1. **La question.** Chaque étudiant reçoit ou choisit une question clinique fictive et la rédige au format PICO, dans un document nommé selon la règle du TP 1.
2. **Les descripteurs MeSH.** Pour chaque élément PICO, on cherche le descripteur dans MeSH Browser. On note le descripteur, deux synonymes (*entry terms*) et sa position dans l'arborescence.
3. **La recherche par concepts.** Dans PubMed, on recherche séparément chaque concept, sous la forme descripteur MeSH OR termes en [tiab]. On note le nombre de résultats.

4. **La combinaison.** Dans Advanced > History, on combine les concepts avec AND, en utilisant les numéros des ensembles (par exemple `#1 AND #2 AND #3`). On note le nombre de résultats et la date.
5. **Les filtres.** On applique successivement le filtre des revues systématiques et des méta-analyses, puis celui des essais randomisés, puis l'intervalle des 5 dernières années. On note l'effet de chaque filtre.
6. **L'enregistrement et l'export.** L'étudiant qui dispose d'un compte My NCBI enregistre la recherche. Tous les étudiants exportent les 10 premiers résultats pertinents au moyen de Send to > Citation manager et importent le fichier `.nbib` dans un logiciel de gestion des références ou, à défaut, ouvrent le fichier dans un éditeur de texte et y repèrent les champs PMID, TI et AU.
7. **Un article clé.** On choisit parmi les résultats une revue systématique récente. On note le type d'étude, l'année, le nombre d'études incluses et la conclusion principale, lue dans le résumé.
8. **Wikipédia.** On ouvre l'article consacré à la maladie de la question en roumain et en anglais. On compare la longueur, la date de la dernière modification et le nombre de références. On vérifie si deux références de l'article existent dans PubMed.
9. **Une page destinée aux patients.** On choisit la première page en roumain renvoyée par un moteur de recherche généraliste pour la maladie en question et on lui applique la liste de vérification du tableau L2.1.
10. **Le chatbot.** On soumet au chatbot la question PICO, sans aucune donnée réelle de patient, en lui demandant explicitement d'indiquer ses sources. On copie intégralement la réponse et l'on note le modèle, la date et le texte de la question [542].
11. **La vérification par rapport au guide.** On compare la recommandation du chatbot avec la conclusion de la revue systématique retenue à l'étape 7 et avec un guide de pratique indiqué par l'enseignant. Chaque référence fournie par le chatbot est recherchée dans PubMed ou d'après son DOI et classée : existante et correcte, existante mais mal décrite, inexistante.
12. **La synthèse.** L'étudiant rédige un paragraphe de 150 mots au plus sur la source qu'il recommanderait à un patient et celle qu'il utiliserait dans un exposé, avec des arguments tirés de la liste de vérification.

Travail à rendre

On remet un seul document (Google Docs ou PDF) qui contient : la question PICO, le tableau des descripteurs MeSH, la requête finale copiée telle quelle depuis PubMed, le nombre de résultats à chaque étape, avec la date, le fichier `.nbib` ou la liste des 10 citations, la liste de vérification remplie pour la page web et pour le chatbot, le résultat de la vérification des références et le paragraphe de synthèse. Barème indicatif, 10 points :

- question PICO claire et descripteurs MeSH corrects (2 points).
- requête syntaxiquement correcte, avec OR à l'intérieur des concepts et AND entre eux (2 points).
- filtres appliqués et documentés, export réalisé (1 point).
- évaluation de Wikipédia et de la page web à l'aide de la liste de vérification (2 points).
- réponse du chatbot documentée, comparée au guide, références vérifiées une à une (2 points).
- synthèse argumentée (1 point).

À l'examen pratique, on peut demander de construire une requête pour une nouvelle question et de classer trois références fournies.

Variante hors ligne

Sans accès à Internet, la séance se déroule sur des supports préparés par l'enseignant. Les étudiants reçoivent, sur papier ou sur clé USB, une question PICO, un extrait de MeSH avec des descripteurs et des synonymes et une liste de 20 citations exportées au préalable depuis PubMed, avec le type de publication. Ils construisent la requête sur papier et déterminent quelles citations seraient retenues par chaque filtre. Pour l'évaluation, ils reçoivent imprimés un article de Wikipédia, une page destinée aux patients et une réponse de chatbot enregistrée au préalable, ainsi que le guide de référence, et ils appliquent la liste de vérification du tableau L2.1. La vérification des références se fait ultérieurement, lorsque le réseau est disponible.

Travaux pratiques 3. Formulaires électroniques de collecte des données : Google Forms et Epi Info

Objectif

L'étudiant conçoit, teste et utilise un formulaire électronique de collecte des données pour une étude fictive, puis transforme les réponses en un tableau descriptif. À la fin de la séance, il doit être capable :

- de rédiger le dictionnaire de données d'une petite étude, avec le type, l'unité et les valeurs autorisées de chaque variable
- de configurer dans Google Forms des règles de validation et l'embranchement du questionnaire en fonction de la réponse
- d'organiser un test pilote et de corriger le formulaire à partir de ses résultats
- d'exporter les réponses dans Google Sheets, de nettoyer les données et de documenter chaque correction
- de construire un tableau descriptif simple à l'aide d'un tableau croisé dynamique
- d'expliquer pourquoi le formulaire ne contient ni nom, ni CNP (code numérique personnel roumain), ni adresse e-mail

La séance applique les notions du chapitre 3 (le dictionnaire de données, la validation à la saisie, la qualité des données, les outils de collecte électronique). La partie consacrée au traitement dans le tableur prolonge la tradition des applications pratiques de la discipline : le protocole de [6] est repris et mis à jour pour Google Sheets 2026, et les exercices dont il est issu figurent dans le volume de 2010 [5]. La partie nouvelle est la collecte des données directement au moyen d'un formulaire électronique, avec des règles de validation fixées avant le recueil.

L'étude fictive. "Le dépistage tensionnel chez les étudiants de première année" est un scénario didactique. Les données sont produites par un jeu de rôle entre camarades (les valeurs sont inventées dans des limites plausibles, ou bien on utilise des mesures faites dans la salle de travaux pratiques, sans lien avec l'identité de la personne). Aucune valeur n'est interprétée sur le plan clinique. Pour le contexte : l'OMS définit l'hypertension comme une pression systolique d'au moins 140 mmHg et/ou une pression diastolique d'au moins 90 mmHg, mesurées deux jours différents [543]. Une seule séance de dépistage ne peut donc signaler qu'une "valeur élevée au dépistage", et non un diagnostic.

Outils

Tous les outils sont gratuits. Les adresses ont été vérifiées le 27 septembre 2026.

- **Google Forms** (https://forms.google.com), application web, version en vigueur à la date de consultation. Un compte Google est nécessaire. La documentation officielle décrit les règles de validation des réponses [123], les sections et le saut en fonction de la réponse [124],

l'association du formulaire à une feuille de calcul [125] et le téléchargement des réponses au format CSV [128].

- **Google Sheets** (https://sheets.google.com), application web, pour le nettoyage des données et le tableau croisé dynamique [544]. Microsoft Excel ou LibreOffice Calc peuvent remplacer Google Sheets pour le fichier CSV.
- **Epi Info 7 pour Windows** (https://www.cdc.gov/epiinfo/), variante de remplacement, sans compte en ligne. Les CDC précisent que le développement et l'assistance technique ont été interrompus en septembre 2025, et que la version de bureau restait fonctionnelle sous Windows 11 à cette date [127]. Le logiciel n'est utilisé ici qu'à des fins didactiques.

Déroulement

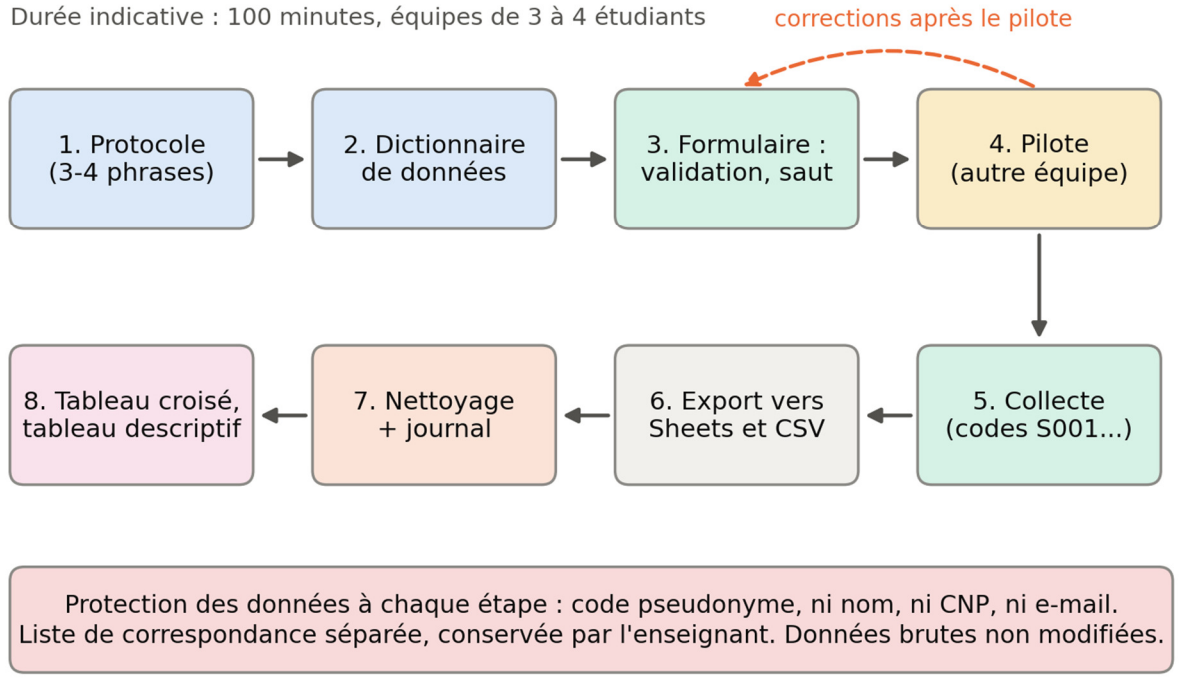


Figure L3.1. Déroulement du TP 3 : du protocole et du dictionnaire de données au formulaire, puis au pilote, à la collecte, à l'export, au nettoyage et au tableau descriptif. La protection des données (pseudonymisation, aucun identifiant direct) s'applique à chaque étape. Schéma original.

La figure L3.1 résume les étapes. La durée indicative est de 100 minutes : 25 pour le protocole et le dictionnaire de données, 30 pour le formulaire et le pilote, 15 pour la collecte, 30 pour l'export, le nettoyage et l'analyse. Les étudiants travaillent en équipes de 3 à 4.

L3.1 Le protocole et le dictionnaire de données

1. L'équipe rédige, en 3 ou 4 phrases, le protocole de l'étude fictive : la population (étudiants de première année âgés de 18 ans ou plus), la variable principale (la pression artérielle), les variables secondaires (âge, sexe, taille, poids, tabagisme, activité physique, antécédents familiaux d'hypertension) et la règle de calcul. Le protocole fictif du TP prévoit trois mesures en position assise, à 1-2 minutes d'intervalle, et l'utilisation de la moyenne des deux dernières.

2. Chaque participant reçoit un code pseudonyme de la forme S001, S002, et ainsi de suite, tiré d'une liste distincte conservée par l'enseignant. La liste de correspondance n'est jamais saisie dans le formulaire.
3. L'équipe complète le dictionnaire de données sur le modèle du tableau L3.1. Pour chaque variable, on fixe le nom court (sans espaces ni signes diacritiques, pour la compatibilité avec les logiciels statistiques), le libellé complet, le type, l'unité, les valeurs autorisées et le code de valeur manquante.

Tableau L3.1. Extrait du dictionnaire de données de l'étude fictive et règle correspondante dans Google Forms. Données fictives.

Variable	Libellé et unité	Type	Valeurs autorisées	Règle dans Google Forms
code	Code pseudonyme du participant	Texte	S suivi de 3 chiffres	Expression régulière : ^S[0-9]{3}$
age	Âge (années révolues)	Numérique entier	18-40	Nombre compris entre 18 et 40
sexe	Sexe	Nominal	F, M, non précisé	Choix multiples
fumeur	Fume actuellement	Nominal	oui, non	Choix multiples avec saut
pas2, pas3	PA systolique, mesures 2 et 3 (mmHg)	Numérique entier	70-250	Nombre compris entre 70 et 250
pad2, pad3	PA diastolique, mesures 2 et 3 (mmHg)	Numérique entier	40-150	Nombre compris entre 40 et 150
act_phys	Jours avec 30 min d'activité physique par semaine	Ordinal	0-7	Liste déroulante 0-7

Les intervalles du tableau sont des règles de plausibilité propres à l'exercice, et non des seuils cliniques. Ils doivent être assez larges pour ne pas rejeter des valeurs réelles inhabituelles, et assez étroits pour intercepter les erreurs de frappe (par exemple, 1280 au lieu de 128).

L3.2 Construction du formulaire dans Google Forms

1. On crée un nouveau formulaire portant le titre de l'étude et une courte description : le but didactique, le fait que les données sont fictives et qu'aucune donnée d'identification n'est demandée.
2. Dans les paramètres, section Réponses, la collecte des adresses e-mail reste désactivée. La documentation indique que le formulaire peut collecter des adresses vérifiées ou saisies par le répondant [128], options qui n'ont pas leur place dans une étude anonyme.
3. La première question est une case à cocher pour le consentement à participer à l'exercice. Elle est marquée comme obligatoire.
4. On ajoute les questions dans l'ordre du dictionnaire de données. Pour chaque question à réponse courte numérique, on active la validation de la réponse, avec la règle "Number" puis "Between" (libellés de l'interface en anglais) [123]. Le message d'erreur personnalisé indique précisément ce qui est attendu ("Saisissez la PA systolique en mmHg, entre 70 et 250").
5. Pour le code pseudonyme, on utilise la validation par expression régulière [123]. L'expression `^S[0-9]{3}$` n'accepte que la lettre S suivie d'exactement trois chiffres.

6. **L'embranchement.** La question "Fumez-vous actuellement ?" est placée en fin de section. Dans le menu de la question, on choisit le saut de section en fonction de la réponse ("Go to section based on answer" dans l'interface en anglais). La réponse "oui" renvoie à une section consacrée au nombre de cigarettes par jour, tandis que la réponse "non" mène directement à la section des mesures [124]. L'embranchement ne fonctionne que pour les questions à choix multiples et à liste déroulante [124].
7. Les variables nominales sont saisies sous forme de choix multiples, et non de texte libre. Ainsi, "F", "f", "fém." et "féminin" ne peuvent plus coexister dans le même fichier.
8. Toutes les questions qui correspondent à la variable principale sont marquées comme obligatoires.

L3.3 Le test pilote

1. Chaque équipe envoie le lien du formulaire à une autre équipe, qui le remplit 3 à 5 fois, y compris avec des valeurs volontairement erronées (un âge de 17 ans, une pression de 1280, un code "s1").
2. L'équipe pilote note la durée de remplissage, les questions ambiguës et les erreurs que le formulaire n'a pas bloquées.
3. Les auteurs corrigent le formulaire. Si une règle de validation a rejeté une valeur correcte, l'intervalle est élargi et la modification est consignée dans le dictionnaire de données.
4. Les réponses du pilote sont supprimées avant la collecte proprement dite. On conserve une courte note décrivant les modifications apportées.

L3.4 La collecte et l'export

1. Chaque étudiant remplit le formulaire pour 3 à 5 "participants" fictifs, en utilisant les codes reçus.
2. Dans l'onglet Réponses, le formulaire est associé à une nouvelle feuille Google Sheets. La feuille reçoit automatiquement les réponses sous forme de tableau [125].
3. À la fin, on désactive l'acceptation des réponses et on télécharge une copie CSV de l'ensemble des réponses, qui sert d'archive des données brutes [128]. Le fichier brut n'est plus modifié.

L3.5 Le nettoyage des données dans Google Sheets

Le nettoyage se fait sur une copie de la feuille de réponses, jamais sur les données brutes. Chaque correction est consignée dans une feuille distincte, "Journal", avec le code du participant, la variable, l'ancienne valeur, la nouvelle valeur et le motif.

1. On renomme les colonnes avec les noms courts du dictionnaire de données (code, age, sexe, et ainsi de suite).
2. **Doublons.** Dans une colonne auxiliaire, on compte le nombre d'occurrences de chaque code : `=COUNTIF(B:B, B2)`. Toute valeur supérieure à 1 signale un double remplissage, qui est examiné (en règle générale, on conserve le premier enregistrement complet).
3. **Valeurs manquantes.** Pour chaque colonne, on calcule le nombre de cellules vides avec `COUNTBLANK`. La proportion de valeurs manquantes figure dans le rapport.
4. **Valeurs hors intervalle.** Bien que le formulaire ait validé les intervalles, on vérifie le minimum et le maximum de chaque variable numérique avec `MIN` et `MAX`.

5. **Cohérence entre champs.** On vérifie que la pression diastolique est inférieure à la pression systolique pour chaque mesure : `=IF(G2<F2, "ok", "à vérifier")`.
6. **Variables dérivées.** On calcule la moyenne des mesures 2 et 3 (`=AVERAGE(F2, H2)` pour la systolique, et de même pour la diastolique), l'indice de masse corporelle (`=poids/(taille/100)^2`) et la catégorie de dépistage : `=IF(OR(pas_m>=140, pad_m>=90), "élevée au dépistage", "sous le seuil")`.

Avec les paramètres régionaux de la Roumanie, comme avec ceux de la France, Google Sheets utilise la virgule comme séparateur décimal et sépare les arguments des fonctions par un point-virgule. Les formules ci-dessus s'écrivent alors, par exemple, `=COUNTIF(B:B; B2)`, ou `=NB.SI(B:B; B2)` avec les noms français des fonctions (SI pour IF, OU pour OR, MOYENNE pour AVERAGE, NB.VIDE pour COUNTBLANK). Les noms anglais des fonctions fonctionnent aussi dans Google Sheets.

L3.6 Le tableau descriptif

1. On sélectionne les données nettoyées et on insère un tableau croisé dynamique [544]. En lignes, on place la variable `sexe`, en colonnes la `catégorie de dépistage`, et en valeurs le nombre de participants (COUNTA sur le code).
2. Un second tableau croisé dynamique affiche la moyenne et l'écart-type de la pression systolique moyenne par sexe.
3. Le résultat est recopié dans un tableau final, sur le modèle du tableau L3.2, avec l'effectif total des participants (N) dans le titre et la mention "données fictives".

Tableau L3.2. Modèle de tableau descriptif pour le rapport. Les cellules sont remplies avec les résultats de l'équipe. Données fictives.

Caractéristique	Féminin	Masculin	Total
Participants, n			
Âge, moyenne (ET)			
PA systolique moyenne, mmHg, moyenne (ET)			
PA diastolique moyenne, mmHg, moyenne (ET)			
Valeur élevée au dépistage, n (%)			
Fumeurs, n (%)			

La discussion en équipe part de trois questions. Quelles erreurs seraient entrées dans les données sans les règles de validation ? Combien de corrections le journal a-t-il enregistrées ? Quelle variable manque au dictionnaire alors qu'elle aurait été nécessaire pour interpréter les résultats (par exemple, la taille du brassard du tensiomètre) ?

L3.7 La variante Epi Info 7

Les équipes qui travaillent sur un ordinateur sous Windows peuvent refaire le formulaire dans Epi Info 7. La page des CDC énumère les composants du logiciel : Form Designer pour le formulaire, Enter Data pour la saisie des données, Classic Analysis et Visual Dashboard pour l'analyse, ainsi que StatCalc [127].

1. Dans Form Designer, on crée un nouveau projet et un formulaire comportant les mêmes variables que le dictionnaire de données.
2. Dans l'éditeur Check Code, on définit les règles d'intervalle, les champs obligatoires et le saut de la question sur le nombre de cigarettes pour les non-fumeurs.
3. On saisit 5 enregistrements fictifs dans Enter Data, dont deux avec des erreurs volontaires, afin de vérifier les règles.
4. Dans Classic Analysis ou Visual Dashboard, on obtient le tableau des fréquences par sexe.
5. L'équipe compare les deux outils selon les critères du tableau 3.3 (coût, validation, piste d'audit, export, assistance à long terme).

Comme le logiciel n'est plus développé, les résultats obtenus avec Epi Info ne sont pas utilisés en dehors de l'exercice.

L3.8 La protection des données

Même dans un exercice, les règles s'appliquent comme dans une étude réelle. Les données concernant la santé sont des données à caractère personnel qui révèlent des informations sur l'état de santé d'une personne, et la pseudonymisation exige que les informations permettant la réidentification soient conservées séparément [118]. Le traitement du code numérique personnel est soumis à des conditions supplémentaires dans la législation roumaine [176]. Dans la pratique du TP, cela signifie que :

- le formulaire ne contient ni nom, ni CNP, ni adresse e-mail, ni numéro de téléphone, ni date de naissance complète
- la liste qui relie les codes aux personnes est conservée par l'enseignant, séparément des données
- la feuille de calcul n'est partagée qu'avec les membres de l'équipe et avec l'enseignant
- à la fin, le formulaire est fermé aux réponses et la feuille n'est conservée que jusqu'à l'évaluation

Pour une étude réelle portant sur des données de patients, Google Forms n'est pas l'outil approprié. On utilise un système EDC institutionnel, doté d'une piste d'audit, avec l'approbation du comité d'éthique (chapitre 3).

Travail à rendre

Chaque équipe remet un seul document (PDF ou lien accessible à l'enseignant), qui contient :

1. le protocole en 3-4 phrases et le dictionnaire de données complet (2 points)
2. le lien du formulaire et des captures d'écran des règles de validation et de l'embranchement, sans données personnelles des utilisateurs (2 points)
3. la note sur le test pilote, avec au moins deux modifications apportées (1 point)
4. la feuille "Journal" avec les corrections de nettoyage (2 points)
5. le tableau descriptif rempli, avec N dans le titre et la mention "données fictives" (2 points)
6. un paragraphe de 5 à 8 phrases qui compare Google Forms à Epi Info 7 ou à REDCap pour une étude clinique réelle (1 point)

Le barème est indicatif, sur 10 points. À l'examen pratique, l'étudiant peut être invité à ajouter une nouvelle règle de validation ou à expliquer une correction du journal.

Variante hors ligne

En l'absence de connexion à Internet, le formulaire est simulé dans une feuille de calcul locale (Microsoft Excel ou LibreOffice Calc). La première feuille contient le dictionnaire de données. Sur la deuxième, chaque colonne reçoit une règle de validation des données (liste de valeurs ou intervalle numérique), équivalente aux règles de Google Forms. L'embranchement est remplacé par une colonne calculée qui signale les réponses incompatibles (par exemple, un nombre de cigarettes renseigné pour un non-fumeur). Epi Info 7, installé au préalable, fonctionne lui aussi sans réseau. Les étapes de nettoyage, le journal et le tableau croisé dynamique restent identiques.

Travaux pratiques 4. Codage des diagnostics avec les outils en ligne de l'OMS (CIM-10 et CIM-11)

Objectif

À la fin de la séance, l'étudiant sait trouver et vérifier un code CIM-10 dans le navigateur de l'OMS, construire un code ou un cluster CIM-11 avec le Coding Tool et le navigateur CIM-11, ordonner correctement les codes d'un cluster pour la morbidité et pour la mortalité et expliquer pourquoi une table de correspondance CIM-10-CIM-11 ne remplace pas le codage manuel. Tous les cas de la séance sont fictifs.

Outils

Tous les outils sont gratuits et fonctionnent dans le navigateur web. Les adresses ont été vérifiées le 27 septembre 2026.

- **Navigateur CIM-10 de l'OMS, version 2019** : https://icd.who.int/browse10/2019/en [133]. Il contient la liste tabulaire et les notes d'inclusion et d'exclusion.
- **Navigateur CIM-11 MMS de l'OMS, version 2026-01** : https://icd.who.int/browse/2026-01/mms/en [141]. La page de chaque entité affiche la description, les exclusions et la section "Postcoordination", avec les axes autorisés pour le code concerné.
- **ICD-11 Coding Tool, MMS 2026-01** : https://icd.who.int/ct [545]. Dans la CIM-11, l'outil de codage remplace l'index imprimé [129].
- **Guide de référence de la CIM-11** (*ICD-11 Reference Guide*), pour les règles de postcoordination et d'ordre des codes [134].
- **Tables de correspondance CIM-10-CIM-11**, disponibles dans l'onglet "Info" du navigateur CIM-11 [135].
- Un tableur (Google Sheets ou Excel) pour la fiche de réponses.

Aucun compte n'est nécessaire. L'ICD API exige une inscription, mais elle n'est pas utilisée dans cette séance [546].

Déroulement

Durée : 100 minutes. L'étudiant travaille individuellement et remplit la fiche de réponses (tableau L4.1). La figure L4.1 montre les étapes qui se répètent pour chaque cas.

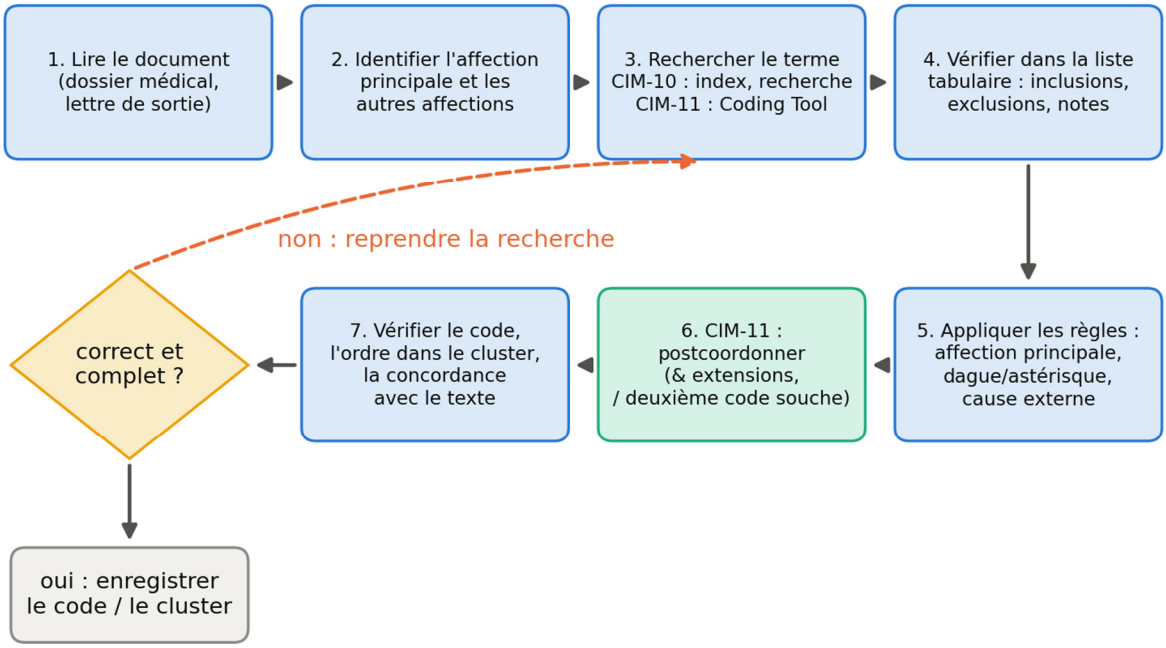


Figure L4.1. Circuit de travail pour le codage d'un diagnostic en CIM-10 et en CIM-11. Schéma d'après les règles de l'OMS [133,134].

L4.1 Préparation (5 minutes)

1. Ouvrez trois onglets du navigateur web : le navigateur CIM-10 2019, le navigateur CIM-11 2026-01 et le Coding Tool.
2. Vérifiez que la version 2026-01 apparaît dans l'en-tête du navigateur CIM-11 et dans celui du Coding Tool. La langue de l'interface reste l'anglais, car les libellés officiels des codes se vérifient dans cette version linguistique. Les codes changent d'une version à l'autre, et les résultats sont toujours rapportés avec la version.
3. Créez une feuille de calcul avec les colonnes du tableau L4.1.

Tableau L4.1. Fiche de réponses (à remplir pour chaque cas).

N°	Affection principale (texte)	Code CIM-10	Code ou cluster CIM-11	Lieu de vérification
1	...	...	...	liste tabulaire / page de l'entité

L4.2 La liste tabulaire de la CIM-10 (15 minutes)

1. Dans le navigateur CIM-10, ouvrez le chapitre IX, "Diseases of the circulatory system" (I00-I99), puis le bloc I20-I25 et la catégorie I21.
2. Lisez la note placée au début du bloc I20-I25 : pour la morbidité, la durée mentionnée dans les catégories I21, I22, I24 et I25 se compte depuis le début de l'épisode ischémique jusqu'à l'admission [133]. Notez dans la fiche ce que signifie "aigu" pour I21 (les inclusions précisent une durée de 4 semaines au plus, soit 28 jours, depuis le début).

3. Ouvrez la catégorie J13 et lisez les exclusions. Notez où se code la pneumonie congénitale à *S. pneumoniae*.
4. Ouvrez G63.2 et observez l'astérisque. Le code de la manifestation (*) s'utilise avec le code de la maladie sous-jacente marqué d'une dague (†), ici E11.4 pour le diabète de type 2 [133].
5. Rappelez-vous la règle de l'OMS : le code trouvé par la recherche (ou dans l'index alphabétique, volume 3) se vérifie toujours dans la liste tabulaire, avec les notes d'inclusion et d'exclusion [133].

L4.3 Le navigateur CIM-11 et le Coding Tool (15 minutes)

1. Dans le Coding Tool, tapez "fracture of neck of femur". La liste de gauche (*Word list*) propose des mots, et la liste de droite (*Destination Entities*) affiche les codes, classés par défaut selon le score de correspondance. Le lien "[Details]" ouvre la description de l'entité.
2. Recherchez le même terme dans le navigateur CIM-11 et ouvrez NC72.20, "Fracture of neck of femur, subcapital". Dans la section "Postcoordination", vous trouverez les axes autorisés : la latéralité (XK8G gauche, XK9K droite), le sous-type de fracture, la fracture ouverte ou fermée (XJ7YM, XJ44E), l'extension à l'articulation et la cause externe associée [141].
3. Recherchez "acute cystitis". Observez que le résultat n'est pas un code unique, mais un cluster déjà assemblé. Notez-le et identifiez la partie qui constitue le code souche et celle qui constitue l'extension.
4. Recherchez "COVID-19 with pneumonia". Observez le séparateur entre les deux codes souches.

Retenez la syntaxe du guide de l'OMS : "&" relie un code souche à ses codes d'extension, "/" relie deux codes souches, et un code d'extension ne peut jamais occuper la première position [134].

L4.4 Codage de dix cas (35 minutes)

L'affection principale est l'affection diagnostiquée à la fin de l'épisode de soins, principalement responsable du besoin de traitement ou d'examens du patient [133]. La CIM-11 conserve cette idée : le motif de l'hospitalisation, établi à la fin de l'épisode [134]. Pour chaque cas, notez aussi dans la fiche le terme principal utilisé pour la recherche, en général le substantif qui désigne la maladie ("infarction", "fracture", "cystitis"). Si le premier terme ne donne pas de résultat, essayez un synonyme.

Pour chaque cas, identifiez l'affection principale, codez-la en CIM-10 (version 2019) et en CIM-11 (MMS 2026-01) et remplissez la fiche. Lorsque le texte demande davantage de détails, utilisez la postcoordination. Les corrigés se trouvent à la fin de la séance, à l'intention de l'enseignant.

1. Homme, 58 ans, hospitalisé pour une douleur thoracique évoluant depuis 2 heures. ECG : sus-décalage du segment ST dans les dérivations antérieures. Diagnostic de sortie : infarctus aigu du myocarde avec sus-décalage du segment ST, paroi antérieure.
2. Femme, 64 ans, consultation de suivi pour un diabète sucré de type 2, sans complication documentée.
3. Homme, 71 ans, pneumonie, avec *Streptococcus pneumoniae* identifié à l'hémoculture.
4. Femme, 55 ans, hypertension artérielle essentielle, sans atteinte d'organe documentée.

5. Femme, 82 ans, a trébuché sur un tapis à son domicile et est tombée. Diagnostic : fracture sous-capitale du col du fémur gauche, fermée.
6. Homme, 19 ans, opéré pour une appendicite aiguë avec péritonite localisée.
7. Femme, 27 ans, asthme allergique, actuellement sans exacerbation.
8. Femme, 30 ans, cystite aiguë.
9. Homme, 66 ans, COVID-19 confirmée par un test de laboratoire, avec pneumonie.
10. Femme, 59 ans, atteinte de diabète sucré de type 2, hospitalisée pour l'exploration et le traitement d'une polyneuropathie diabétique, qui constitue l'affection principale de l'épisode.

Pour le cas 10, comparez les deux solutions. La CIM-10 utilise le système dague-astérisque, la CIM-11 un cluster à deux codes souches. Pour la morbidité, le guide de l'OMS place en première position la manifestation qui répond à la définition de l'affection principale, suivie du code du diabète, comme dans l'exemple de la rétinopathie diabétique 9B71.0Z/5A11 [134].

L4.5 Exercice de postcoordination (10 minutes)

En partant du cas 1, ajoutez l'information "lésion de l'artère coronaire descendante antérieure gauche". Recherchez l'extension d'anatomie dans le navigateur CIM-11 et construisez le cluster complet. Répondez ensuite : quelle information du cluster ne peut pas être exprimée en CIM-10 ?

En partant du cas 5, construisez un cluster qui contienne, dans l'ordre, la fracture, la latéralité, le caractère fermé de la fracture et la cause externe. Le guide de l'OMS donne un exemple de même structure pour une fracture du cubitus gauche après une chute : NC32.2&XK8G&XJ7YM/PA60&XE1DA&XE53A [134]. Expliquez pourquoi PA60 est précédé de "/" et non de "&".

L4.6 La correspondance CIM-10-CIM-11 (10 minutes)

1. Depuis l'onglet "Info" du navigateur CIM-11, ouvrez les tables de correspondance CIM-10-CIM-11.
2. Recherchez dans la table les codes CIM-10 que vous avez attribués aux cas 1, 5 et 8. Notez dans la fiche ce que propose la table.
3. Comparez la proposition de la table avec le code ou le cluster que vous avez construit. En quoi diffèrent-ils ? Quels détails du texte clinique ne peuvent pas être récupérés à partir du seul code CIM-10 ?

L'OMS précise que les tables de correspondance servent à comparer les données entre révisions, et non à la conversion automatique [135]. Sur des données réelles, les tables ont bien fonctionné pour les causes de décès : appliquées à 24 645 décès survenus en Alberta, elles ont permis de mettre en correspondance 96,6 % des codes CIM-10 chez l'adulte [144]. Pour la morbidité, en revanche, l'information absente du code CIM-10 reste absente après la mise en correspondance. L'écart de granularité s'observe aussi en sens inverse : dans la correspondance canadienne de la CIM-11 vers la CIM-10-CA, 65 % des codes CIM-10-CA cibles étaient plus larges que le code CIM-11 source [145].

L4.7 La cause initiale du décès et l'ordre dans le cluster (10 minutes)

1. Un certificat de décès fictif indique : (a) pneumonie, due à (b) une fracture de la hanche, due à (c) un trébuchement sur un tapis. Identifiez le point de départ de la séquence, puis son code

en CIM-10 et en CIM-11. Le guide de l'OMS traite cet exemple et précise que le point de départ est l'événement externe [134].

2. Un patient atteint de diabète sucré de type 2 est hospitalisé dans un coma diabétique. Écrivez le cluster CIM-11 pour la morbidité. Le même patient décède du coma. Écrivez le cluster pour la mortalité. Expliquez pourquoi l'ordre s'inverse.

Pièges fréquents

- Copier le premier résultat de la recherche sans ouvrir la page de l'entité. Les exclusions renvoient souvent à un autre code.
- S'arrêter à la catégorie à trois caractères alors qu'il existe des sous-catégories (par exemple J45 au lieu de J45.0).
- Placer le code de cause externe comme affection principale en morbidité. En morbidité, la lésion traumatique est l'affection principale, et la cause externe s'y ajoute.
- Écrire un code d'extension CIM-11 seul ou au début du cluster.
- Utiliser le séparateur "&" entre deux codes souches.
- Rapporter un code sans sa version. Un même concept peut avoir un autre code dans une autre version.
- Accepter un code proposé par un chatbot sans vérification. Dans une étude de 2024, les grands modèles de langage, à qui l'on fournissait la description officielle du code, ont généré le code exact dans moins de la moitié des cas [165].

Travail à rendre

Le fichier de la feuille de calcul (ou la fiche imprimée), avec :

- le tableau L4.1 rempli pour les 10 cas : 6 points (0,3 point pour le code CIM-10 correct et complet, 0,3 point pour le code ou le cluster CIM-11 correct, avec l'ordre et les séparateurs corrects)
- les clusters de l'exercice de postcoordination et la réponse aux questions : 1 point
- la comparaison avec la table de correspondance pour trois cas : 1 point
- l'exercice de mortalité et l'explication de l'ordre dans le cluster : 1 point
- d'office : 1 point.

Barème indicatif : un code correct, mais assorti d'une version erronée ou non vérifié dans la liste tabulaire, reçoit la moitié des points. À titre de comparaison, lors de l'essai sur le terrain canadien, le temps moyen de codage avec la CIM-11 est passé de 23,6 à 9,9 minutes par dossier à mesure que les codeurs gagnaient en expérience [137], et en Corée, des spécialistes expérimentés ont obtenu une exactitude moyenne de 71,6 % pour le codage en CIM-11 de termes diagnostiques isolés [138]. Les erreurs sur les premiers cas sont normales. L'objectif est que l'étudiant soit capable de les expliquer.

Variante hors ligne

Si la salle n'a pas de connexion Internet, l'enseignant prépare à l'avance, sur papier ou sous forme de fichiers PDF, des extraits de la liste tabulaire de la CIM-10 (chapitres IV, VI, IX, X, XI, XIV, XIX, XX et XXII, uniquement les catégories utilisées dans les cas) et des captures des pages CIM-11 des entités figurant dans les corrigés, y compris la section "Postcoordination". Les étudiants travaillent sur les mêmes cas, et la recherche est remplacée par le parcours de la hiérarchie dans les extraits. Autre possibilité, l'OMS permet l'installation locale du service ICD

API, par exemple dans un conteneur Docker, sur un ordinateur du laboratoire configuré à l'avance [546]. L'exercice L4.6 est reporté à une séance avec réseau ou réalisé sur un extrait imprimé de la table de correspondance.

Corrigés (pour l'enseignant)

Tous les codes ont été vérifiés le 27 septembre 2026 dans le navigateur CIM-10 de l'OMS, version 2019 [133], et dans le navigateur CIM-11 MMS, version 2026-01 [141]. L'ordre des codes dans les clusters suit le guide de référence de la CIM-11 [134].

Tableau L4.2. Corrigés des cas 1-10 (morbidité). Source : [133,134,141].

Cas	CIM-10 (2019)	CIM-11 (MMS 2026-01)	Observations
1	I21.0	BA41.0&XA7RE3	XA7RE3 = paroi antérieure du cœur
2	E11.9	5A11	5A11 n'a pas de sous-catégories
3	J13	CA40.07	J13 est un code à trois caractères, sans subdivisions
4	I10	BA00.Z	BA00.Z = hypertension essentielle, sans précision
5	S72.0 ; cause externe W01	NC72.20&XK8G&XJ44E/PA60	XK8G gauche, XJ44E fermée, PA60 chute de plain-pied
6	K35.3	DB10.01	appendicite aiguë avec péritonite localisée
7	J45.0	CA23.02	asthme allergique, non compliqué
8	N30.0	GC00.1&XT5R	GC00.1 cystite infectieuse, XT5R aigu
9	U07.1 (et, facultativement, un code supplémentaire pour la pneumonie)	RA01.0/CA40.1Z	CA40.1Z = pneumonie virale, sans précision
10	E11.4† G63.2*	8C03.0/5A11	en morbidité, la manifestation principale en première position

Réponses aux exercices L4.5 et L4.7 :

- L4.5, cas 1 : BA41.0&XA7RE3&XA7NQ7. L'artère atteinte (XA7NQ7, artère coronaire descendante antérieure gauche) n'a pas d'équivalent en CIM-10, où I21.0 ne décrit que la paroi. Le cluster reproduit l'exemple du guide de l'OMS [134].
- L4.5, cas 5 : NC72.20&XK8G&XJ44E/PA60. PA60 est un code souche du chapitre 23 (causes externes), il se relie donc par "/". Les extensions XK8G et XJ44E appartiennent au code de la fracture et s'écrivent immédiatement après lui.
- L4.7.1 : le point de départ est le trébuchement, c'est-à-dire l'événement externe. Son code est W01 en CIM-10 (chute de plain-pied résultant de glissade, faux-pas et trébuchement) et PA60 en CIM-11 (chute accidentelle de plain-pied ou d'une hauteur de 1 mètre ou moins).
- L4.7.2 : morbidité 5A23/5A11 (le coma est le motif de l'hospitalisation), mortalité 5A11/5A23 (le diabète a déclenché la chaîne d'événements). En CIM-10, les deux situations utilisent le code combiné E11.0, diabète sucré de type 2 avec coma [133,134].
- L4.3 : le cluster pour "acute cystitis" est GC00.1&XT5R, et celui pour "COVID-19 with pneumonia, SARS-CoV-2 identified" est RA01.0/CA40.1Z [141].

Travaux pratiques 5. Citer correctement, prévenir le plagiat et rédiger un essai scientifique

Objectif

À l'aide d'un gestionnaire de références, l'étudiant construit une bibliographie en style Vancouver pour un essai de synthèse de 1 200 à 1 500 mots sur un thème du chapitre 5. Il s'exerce à mettre en forme cinq types de sources (article, livre, chapitre, page web, texte normatif), distingue la paraphrase de la copie, lit un rapport de similitude et rédige la déclaration relative à l'utilisation des outils d'intelligence artificielle (IA) générative. La séance applique les notions de la section 5.6.

Outils

- **PubMed** (https://pubmed.ncbi.nlm.nih.gov/, consulté le 27 septembre 2026), pour la recherche et l'export. Le bouton Cite propose le format NLM, sur lequel repose le style Vancouver, et télécharge le fichier .nbib que les gestionnaires de références importent [41].
- **Zotero 10**, téléchargé gratuitement sur https://www.zotero.org/download/ (consulté le 27 septembre 2026, date à laquelle la version pour Windows était la 10.0.3), avec l'extension Zotero Connector pour Chrome, Firefox ou Edge [547]. Les modules pour Microsoft Word et LibreOffice sont inclus dans le logiciel et s'installent automatiquement au premier démarrage ; le travail dans Google Docs passe par l'extension Zotero Connector, avec Zotero ouvert [548]. Le style Vancouver s'installe au besoin depuis le dépôt de styles, qui contient plus de 10 000 styles au format CSL [549].
- **Solution de remplacement : Mendeley Reference Manager**, version 2.149.0 pour Windows 10 ou ultérieur (https://www.mendeley.com/download-reference-manager/windows, consulté le 27 septembre 2026), avec le module Mendeley Cite pour Word [550].
- **Traitement de texte** : Microsoft Word, LibreOffice Writer ou Google Docs.
- **Le guide de format** de la Bibliothèque nationale de médecine des États-Unis (NLM), avec des exemples pour chaque type de source [201], et les recommandations de l'ICMJE, mises à jour en janvier 2026 [202].
- **Le logiciel antiplagiat** mis à disposition par l'université, par l'intermédiaire du compte institutionnel. Le TP ne dépend d'aucun produit particulier.

Déroulement

La figure L5.1 présente les six étapes du TP. La boucle du bas en est la partie la plus importante : le rapport de similitude ne clôt pas le travail, il renvoie à la rédaction lorsque apparaissent des correspondances injustifiées ou des affirmations sans source.

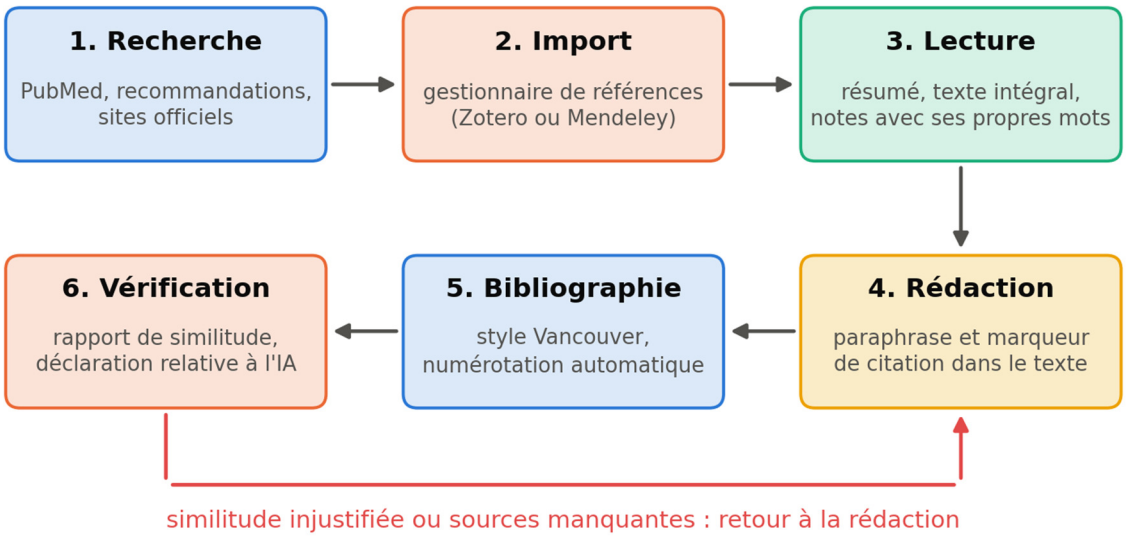


Figure L5.1. Déroulement du TP 5 : de la recherche des sources à la vérification finale. La boucle rouge indique le retour à la rédaction après la lecture du rapport de similitude. Schéma original.

L5.1 Anatomie d'une référence Vancouver

Le style Vancouver comporte deux composantes. Dans le texte, les sources sont signalées par des chiffres arabes, dans l'ordre de leur première apparition : [1], [2], [3]. Une source citée de nouveau conserve son numéro initial. À la fin, la liste bibliographique suit le même ordre numérique, et non l'ordre alphabétique [201,202].

Règles de base pour la liste finale, d'après le guide de la NLM [201] :

1. Les auteurs sont indiqués par leur nom de famille suivi des initiales, sans points entre les initiales (Neprash HT). On mentionne les six premiers auteurs, puis "et al." s'ils sont plus nombreux.
2. Le titre de l'article ne prend de majuscule qu'au début et aux noms propres.
3. Le nom de la revue est abrégé selon le catalogue de la NLM (JAMA Health Forum, N Engl J Med), sans points.
4. Viennent ensuite l'année, le volume, le numéro entre parenthèses et les pages : 2022;3(12):e224873. L'intervalle de pages est abrégé (2129-31 au lieu de 2129-2131).
5. Pour les sources en ligne, on ajoute [Internet], la date de consultation [cited ...] et l'adresse, après "Available from:".

Le tableau L5.1 donne un modèle pour chaque type de source. Tous les exemples sont des références réelles, liées au chapitre 5, et leurs données bibliographiques ont été vérifiées dans PubMed, dans CrossRef ou sur le site de l'émetteur.

Tableau L5.1. Modèles Vancouver et exemples réels. Source : [62,170,176,183,201,202].

Type	Modèle	Exemple
Article	Auteurs. Titre. Revue. Année;Vol(No):pages.	Neprash HT, McGlave CC, Cross DA, Virnig BA, Puskarich MA, Huling JD, et al. Trends in ransomware attacks on US hospitals, clinics, and other health care delivery organizations, 2016-2021. JAMA Health Forum. 2022;3(12):e224873.
Livre	Auteurs ou directeurs. Titre. Édition. Lieu: Éditeur; Année.	Shortliffe EH, Cimino JJ, editors. Biomedical informatics: computer applications in health care and biomedicine. 5th ed. Cham: Springer; 2021.
Chapitre	Auteurs. Titre du chapitre. In: Directeurs, editors. Titre du livre. Édition. Lieu: Éditeur; Année. p. pages.	Goodman KW, Miller RA. Ethics in biomedical and health informatics: users, standards, and outcomes. In: Shortliffe EH, Cimino JJ, editors. Biomedical informatics: computer applications in health care and biomedicine. 5th ed. Cham: Springer; 2021. p. 391-423.
Page web	Auteur. Titre [Internet]. Lieu: Éditeur; date [updated date; cited date]. Available from: URL	International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals [Internet]. [place unknown]: ICMJE; [updated 2026 Jan; cited 2026 Sep 27]. Available from: https://www.icmje.org/recommendations/
Texte normatif	Titre de l'acte. Publication officielle. Date;(numéro).	Legea nr. 190/2018 privind măsuri de punere în aplicare a Regulamentului (UE) 2016/679 al Parlamentului European și al Consiliului din 27 aprilie 2016 privind protecția persoanelor fizice în ceea ce privește prelucrarea datelor cu caracter personal și privind libera circulație a acestor date și de abrogare a Directivei 95/46/CE (Regulamentul general privind protecția datelor). Monitorul Oficial al României, Partea I. 2018 Jul 26;(651).

Les erreurs de format les plus fréquentes chez les étudiants sont prévisibles : une liste alphabétique au lieu d'une numérotation dans l'ordre de citation, les 15 auteurs d'un article au lieu des six premiers suivis de "et al.", le titre de la revue écrit en entier ou abrégé au jugé, l'absence de date de consultation pour les pages web et la citation de la notice PubMed au lieu de l'article. Une autre erreur, plus grave, consiste à citer une source secondaire (un manuel, une encyclopédie en ligne, une réponse d'un chatbot) pour un chiffre issu d'une étude primaire. Le chiffre doit être recherché dans l'article original, et c'est celui-ci que l'on cite.

Deux remarques. Pour un chapitre, on cite les auteurs du chapitre, et non les directeurs de l'ouvrage, et les pages du chapitre sont obligatoires. Pour les textes normatifs, c'est la date de publication au Journal officiel de la Roumanie (Monitorul Oficial) qui compte, et non la date d'adoption : la loi n° 190/2018 a été adoptée le 18 juillet et publiée le 26 juillet 2018 [176].

L5.2 Le gestionnaire de références

Les étapes ci-dessous concernent Zotero. Pour Mendeley, la logique est la même, avec des intitulés de menus différents.

1. On installe Zotero et l'extension Zotero Connector pour le navigateur. On crée une nouvelle collection, nommée "Essai TP5".
2. Dans PubMed, on recherche le thème choisi, par exemple `ransomware[ti] AND hospital`, avec le filtre d'années 2020-2026. Pour chaque article pertinent, on clique sur le bouton Zotero de la barre du navigateur. Autre possibilité : dans Zotero, on utilise le bouton "Add Item by Identifier" (la baguette magique) et on colle le PMID ou le DOI.
3. Pour chaque notice, on vérifie manuellement les noms des auteurs, le titre, la revue, l'année, le volume, les pages et le DOI. L'import automatique se trompe parfois sur les signes diacritiques, sur les auteurs collectifs et sur les pages électroniques (du type e224873).

4. Les pages web et les textes normatifs sont ajoutés manuellement ("New Item" > "Web Page" et "Statute", respectivement), en renseignant la date de consultation.
5. Dans Word apparaît l'onglet Zotero. On choisit le style "NLM/Vancouver: Citing Medicine 2nd edition (citation-sequence)", installé au besoin depuis le dépôt de styles Zotero [549]. La commande "Add/Edit Citation" insère les citations dans le texte, et "Add/Edit Bibliography" génère la liste finale.
6. On vérifie la numérotation : la première source citée dans le texte doit être [1] dans la liste. Si un paragraphe est déplacé, on clique sur "Refresh" et la numérotation est refaite automatiquement.
7. À la fin, la collection est exportée par un clic droit sur la collection > "Export Collection...", dans le format choisi (par exemple RIS). L'export n'est pas une sauvegarde : les articles réimportés perdent le lien avec les citations des documents Word déjà rédigés [551]. Un fichier RIS peut aussi être importé dans Mendeley Reference Manager, par "Add New" > "Import Library" [552].

Les citations insérées par le gestionnaire sont des champs actifs du document Word. Si le document est converti en texte brut ou copié dans un autre logiciel, le lien avec la bibliothèque est perdu et la numérotation ne se met plus à jour. C'est pourquoi on conserve toujours une copie du document dont les champs sont intacts.

L5.3 Paraphrase, citation, copie

Ce qui sépare l'utilisation correcte d'une source du plagiat n'est pas le nombre de mots modifiés, mais le renvoi à la source et la reformulation réelle de l'idée. La loi roumaine n° 199/2023 range explicitement dans le plagiat la reprise des idées, des données et des méthodes sans renvoi à la source, et pas seulement celle du texte [198]. L'exercice porte sur une phrase du résumé de l'article de Neprash et collaborateurs [183] :

- **Texte original** (extrait) : "the annual number of ransomware attacks more than doubled from 43 to 91".
- **Traduction mot à mot** (inacceptable en tant que paraphrase) : "le nombre annuel d'attaques par rançongiciel (*ransomware*) a plus que doublé, passant de 43 à 91 [1]". La source est indiquée, mais la formulation des auteurs originaux apparaît, sans guillemets, comme une formulation personnelle. Que certains logiciels antiplagiat ne détectent pas les traductions ne rend pas le procédé correct.
- **Citation correcte** : Neprash et collaborateurs montrent que "the annual number of ransomware attacks more than doubled from 43 to 91" [1]. La citation directe s'emploie rarement, pour des formulations que l'on ne saurait mieux rendre.
- **Paraphrase correcte** : "Aux États-Unis, la fréquence des attaques par rançongiciel contre les établissements de soins a été multipliée par plus de deux en six ans, entre 2016 et 2021 [1]." L'idée est la même, la structure et les mots sont ceux de l'auteur de l'essai, et la source est indiquée.

L'étudiant rédige ensuite, pour deux sources de son choix, une paraphrase et un résumé de deux phrases pour chacune, avec citation. Une méthode utile : on lit le paragraphe, on ferme la source, on écrit l'idée de mémoire, et c'est seulement ensuite qu'on la compare à l'original, pour vérifier l'exactitude des chiffres.

L5.4 Le rapport de similitude

Le logiciel antiplagiat compare le texte à des bases de documents publiés et à des travaux remis antérieurement. Le résultat est un pourcentage de similitude et une liste de sources, avec les passages colorés. Le pourcentage seul n'indique pas s'il y a plagiat : un passage identique peut être une citation correcte, et l'évaluateur doit analyser chaque correspondance (section 5.6.1). L'étudiant procède ainsi :

1. Il téléverse la version finale de l'essai, bibliographie comprise.
2. Il lit la liste des sources par ordre décroissant de similitude. Si une seule source dépasse quelques pour cent, le passage concerné est vérifié en premier.
3. Il classe chaque correspondance : citation signalée par des guillemets et accompagnée de sa source (acceptable), expression technique usuelle, du type "analyse d'impact relative à la protection des données" (acceptable), entrée de la bibliographie (acceptable) ou phrase reprise sans guillemets (à réécrire).
4. Il réécrit les passages de la dernière catégorie et soumet de nouveau le texte, si le système le permet.
5. Il note dans la fiche de remise le pourcentage initial, le pourcentage final et, brièvement, ce qu'il a modifié.

On ne cherche pas à tromper le logiciel en remplaçant des lettres, en choisissant mécaniquement des synonymes ou en recourant à une traduction automatique aller-retour. Le texte obtenu est difficile à lire, et l'idée reste reprise sans source.

L5.5 Déclarer l'utilisation de l'IA générative

Les outils d'IA générative ne peuvent être utilisés que dans les limites fixées par le titulaire de la discipline. Les recommandations de l'ICMJE demandent de décrire la manière dont ils ont été utilisés, et les outils d'IA ne peuvent pas figurer comme auteurs, car ils ne peuvent pas répondre de l'exactitude et de l'originalité du travail. Le contenu produit par l'IA n'est pas cité comme source primaire [202]. Le COPE demande en outre de nommer l'outil utilisé [203]. Pour l'essai, on ajoute à la fin une courte section, par exemple :

Exemple

- Déclaration relative à l'utilisation de l'IA : "J'ai utilisé [nom de l'outil, version] pour la correction grammaticale de la section Introduction et pour proposer une structure de la section Discussion. Toutes les sources ont été recherchées et lues par moi-même dans PubMed et sur les sites des émetteurs. Je réponds de l'ensemble du contenu."
- Si l'IA n'a pas été utilisée : "Je n'ai pas utilisé d'outils d'IA générative pour rédiger cet essai."

Les références proposées par un chatbot ne sont pas copiées dans la bibliographie avant d'avoir été effectivement retrouvées dans PubMed, dans CrossRef ou sur la page officielle de l'émetteur.

L5.6 La structure de l'essai

L'essai de synthèse suit une structure proche du format IMRaD (Introduction, Méthodes, Résultats et Discussion), adaptée à une revue courte :

1. **Titre** informatif, de 15 mots au plus.
2. **Introduction** (150-200 mots) : le contexte, l'intérêt du thème pour le médecin et la question précise à laquelle répond l'essai.
3. **Méthodes** (100-150 mots) : les bases de données consultées, les termes de recherche, la période (2020-2026), les critères d'inclusion et le nombre de sources retenues.
4. **Résultats** (500-600 mots) : la synthèse par thèmes, et non le résumé successif de chaque article. Cette partie peut inclure un tableau des principales études (auteur, année, pays, N, résultat).
5. **Discussion** (300-400 mots) : la conclusion que soutiennent les sources, les points de désaccord et les limites des données probantes comme de la recherche personnelle.
6. **Conclusion** (2-3 phrases) et **Bibliographie** Vancouver, avec 8 à 12 sources, dont au moins 6 articles de revues indexées dans PubMed.
7. **Déclaration relative à l'IA**.

Thèmes suggérés : l'hypoxémie occulte et l'oxymétrie de pouls, le consentement dynamique pour la recherche, l'impact clinique des attaques par rançongiciel, l'anonymisation des images médicales, les politiques des revues à l'égard de l'IA générative.

Travail à rendre

Sont déposés sur la plateforme du cours, avant la date limite annoncée :

1. L'essai (1 200-1 500 mots, hors bibliographie), au format .docx, avec les citations insérées depuis le gestionnaire de références.
2. L'export de la bibliothèque (fichier .ris ou .bib) contenant toutes les sources citées.
3. La fiche de travail, avec : les cinq références du tableau L5.1 remises en forme pour une source personnelle de chaque type, les deux paraphrases de la section L5.3 et les notes sur le rapport de similitude de la section L5.4.
4. La déclaration relative à l'IA.

Barème indicatif (10 points) : structure IMRaD et clarté de la question, 2 points. Synthèse et exactitude des affirmations par rapport aux sources, 3 points. Bibliographie Vancouver complète, sans erreurs de format et correctement numérotée, 2 points. Paraphrases et analyse du rapport de similitude, 2 points. Déclaration relative à l'IA et respect de la date limite, 1 point. Un passage repris sans source ou une référence inexistante annule les points de la synthèse et fait l'objet d'une discussion avec le titulaire de la discipline.

Variante hors ligne

En l'absence d'accès à Internet, le TP se fait sur papier ou dans un traitement de texte local. L'enseignant distribue des fiches imprimées reproduisant les premières pages de cinq sources (un article, un livre, un chapitre, une page web imprimée, avec la date de consultation notée, et un texte normatif paru au Monitorul Oficial). L'étudiant rédige à la main les références Vancouver, d'après les modèles du tableau L5.1, puis les compare à la version modèle. Zotero fonctionne aussi hors ligne pour les notices déjà enregistrées dans la bibliothèque, si bien que la citation dans Word peut être exercée sur une bibliothèque préparée à l'avance. Les exercices de paraphrase (L5.3) et la structure de l'essai (L5.6) ne nécessitent pas de réseau. La vérification de similitude est reportée jusqu'au rétablissement de la connexion.

Travaux pratiques 6. Analyse des signaux ECG et EEG avec des logiciels libres

Objectif

L'étudiant ouvre des enregistrements physiologiques publics, les visualise avec des logiciels gratuits et en extrait des mesures simples et reproductibles. À la fin de la séance, l'étudiant sera capable :

- d'identifier, sur un tracé ECG annoté, les battements normaux et les battements prématurés.
- de mesurer 20 intervalles RR consécutifs et de calculer la fréquence cardiaque moyenne, le SDNN et le RMSSD dans une feuille de calcul.
- de mettre en évidence le repliement (*aliasing*) par le sous-échantillonnage d'un signal simple.
- de reconnaître le rythme alpha occipital sur un EEG enregistré les yeux fermés et d'expliquer pourquoi il s'atténue à l'ouverture des yeux.

La séance applique les notions du chapitre 6 (échantillonnage, quantification, détection des QRS, variabilité de la fréquence cardiaque, VFC). Toutes les données sont publiques et anonymisées. On ne travaille pas sur des données de patients de la clinique.

Outils

Tous les outils sont gratuits. Les adresses et les versions ont été vérifiées le 27 septembre 2026.

- **PhysioNet** (https://physionet.org/), la ressource ouverte de signaux physiologiques maintenue par le Laboratoire de physiologie computationnelle du MIT [226]. La citation standard demandée aux utilisateurs est l'article de 2026 qui décrit la plateforme [227].
- **MIT-BIH Arrhythmia Database**, version 1.0.0 : 48 extraits d'ECG de 30 minutes chacun, deux canaux, 360 échantillons par seconde, résolution de 11 bits sur 10 mV, annotations vérifiées par au moins deux cardiologues, licence Open Data Commons Attribution [213].
- **EEG Motor Movement/Imagery Dataset**, version 1.0.0 : 109 sujets, 64 électrodes, 160 Hz, fichiers EDF+. Les deux premiers enregistrements de chaque sujet sont des lignes de base d'une minute chacune, les yeux ouverts (R01) et les yeux fermés (R02) [553].
- **LightWAVE 0.71**, le visualiseur de PhysioNet qui fonctionne dans le navigateur, sans installation (https://physionet.org/lightwave/) [228].
- **EDFbrowser 2.15** (https://www.teuniz.net/edfbrowser/), logiciel libre sous licence GPL version 3, pour Windows 64 bits et Linux. Il ouvre les fichiers EDF, EDF+ et BDF, convertit d'autres formats (dont le format MIT), calcule le spectre de puissance par FFT et intègre un détecteur de QRS de type Pan-Tompkins avec export des intervalles RR. L'auteur précise que le logiciel ne doit pas être utilisé dans des applications critiques pour la sécurité du patient [216].
- **Feuille de calcul** : Google Sheets (application web) ou Microsoft Excel.
- **En option, Python 3** avec la bibliothèque numpy et le paquet wfdb 4.3.1, qui lit directement les enregistrements de PhysioNet [554,555].

Déroulement

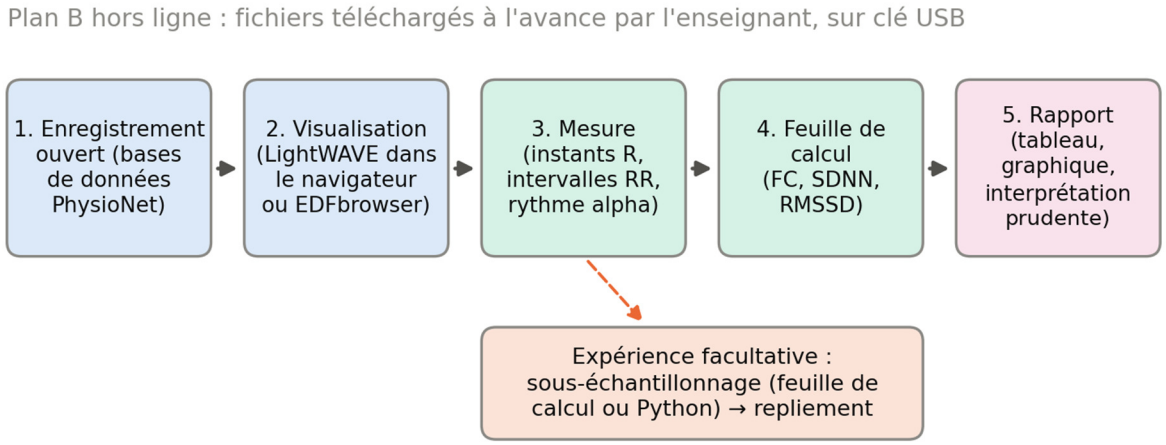


Figure L6.1. Déroulement du TP 6 : de l'enregistrement ouvert au rapport final. Schéma original.

La figure L6.1 résume les étapes. La durée indicative est de 100 minutes : 45 pour l'ECG, 20 pour l'expérience d'échantillonnage et 35 pour l'EEG.

L6.1 Partie A. ECG : visualisation et annotations

1. Ouvrez la page de la base MIT-BIH Arrhythmia (https://physionet.org/content/mitdb/1.0.0/). Notez dans votre cahier la fréquence d'échantillonnage (360 Hz) et la résolution (11 bits sur 10 mV). Calculez le pas de quantification : 10 mV / 2^11 ≈ 4,9 µV.
2. Sur la même page, repérez la section qui indique les publications à citer. Recopiez-les dans votre rapport. L'utilisation d'une base publique ne dispense pas de citer la source.
3. Ouvrez LightWAVE avec la base MIT-BIH : https://physionet.org/lightwave/?db=mitdb/1.0.0. Dans l'onglet **Choose input**, choisissez l'enregistrement **100** dans la liste **Record**.
4. Passez à l'onglet **View/edit**. L'écran affiche 10 secondes de deux dérivations, MLII et V5, avec la grille de fond et les annotations de référence (ligne `atr`). Les battements normaux sont marqués d'un point. Dans les 10 premières secondes apparaît aussi une annotation **A**, c'est-à-dire un battement atrial prématuré (*atrial premature contraction*) [555].
5. Passez à l'onglet **Tables**. Le résumé indique la durée de l'enregistrement (00:30:06), la fréquence de 360 échantillons par seconde et les 2 274 annotations : 2 239 battements normaux (N), 33 battements atriaux prématurés (A), une extrasystole ventriculaire (V) et une annotation de rythme, `(N` [228]. On y voit aussi que chaque signal compte 200 unités de conversion par millivolt, si bien qu'une unité correspond à 5 µV. Comparez avec le pas calculé à l'étape 1.
6. Cochez l'option **View annotations as text** pour obtenir la liste des annotations avec leurs instants. Si la liste ne s'affiche pas, revenez à **View/edit** et lisez les instants des pics R sur l'axe du temps.

Un calcul de contrôle, utile par la suite : les 2 273 annotations de battements (2 239 + 33 + 1) se répartissent sur 30 minutes et 6 secondes, soit 1 806 secondes. La fréquence cardiaque moyenne sur l'ensemble de l'enregistrement est donc d'environ 2 273 × 60 / 1 806 ≈ 75,5 battements par minute. La valeur que vous obtenez sur un segment court doit être du même ordre

de grandeur. Un écart important signale généralement une erreur d'unité (secondes au lieu de millisecondes) ou un battement omis.

L6.2 Partie B. Intervalles RR, fréquence cardiaque et VFC dans la feuille de calcul

1. Choisissez un segment d'environ 20 secondes **sans** annotations A ou V. Notez les instants de 21 pics R consécutifs dans la colonne B du tableau L6.1 (en secondes, avec trois décimales si vous utilisez la liste des annotations).
2. Calculez dans la colonne C les intervalles RR en millisecondes : pour la ligne 3, `=(B3-B2)*1000`. Vous obtenez 20 intervalles.
3. Dans la colonne D, la fréquence cardiaque instantanée : `=60000/C3`.
4. Dans la colonne E, les différences successives : pour la ligne 4, `=C4-C3`. Vous obtenez 19 différences.
5. Calculez, sous le tableau :

- RR moyen : `=AVERAGE(C3:C22)`
- FC moyenne : `=60000/AVERAGE(C3:C22)`
- SDNN : `=STDEV.S(C3:C22)`
- RMSSD : `=SQRT(SUMSQ(E4:E22)/COUNT(E4:E22))`

Dans les interfaces traduites des logiciels, les noms des fonctions peuvent différer (en français, par exemple, MOYENNE pour AVERAGE, RACINE pour SQRT et NB pour COUNT).

6. Construisez le tachogramme : le graphique des intervalles RR (axe vertical) en fonction de l'instant du battement (axe horizontal). La figure L6.2 montre à quoi ressemble un tachogramme et comment un battement prématuré s'y distingue.
7. Refaites le calcul sur un segment qui **contient** un battement A. Comparez le SDNN et le RMSSD avec les valeurs du segment propre et expliquez la différence. Les recommandations relatives à la publication des résultats exigent d'identifier et de traiter les battements ectopiques avant le calcul de la VFC, de préférence par une inspection visuelle [212].

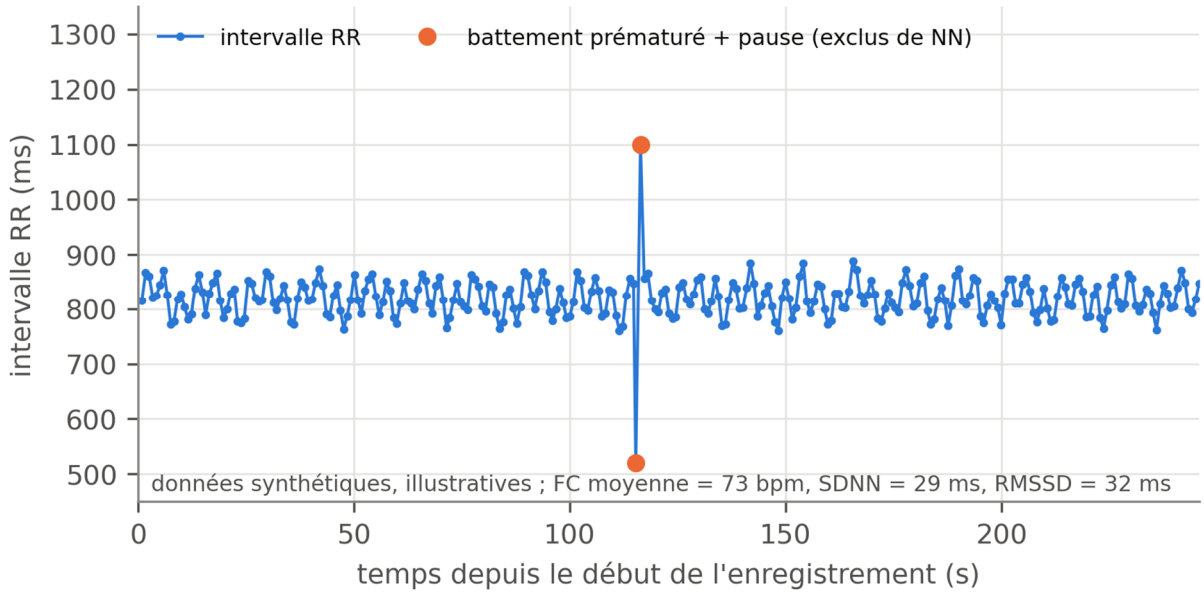


Figure L6.2. Tachogramme RR calculé à partir d'une série synthétique de 300 battements, avec un battement prématuré suivi d'une pause. Les deux intervalles signalés sont exclus de la série NN avant le calcul du SDNN et du RMSSD. Données synthétiques, illustratives, générées par le script fig_L06.py.

Interprétez avec prudence. Avec 20 intervalles, les indicateurs n'ont qu'une valeur didactique. Le rapport de recommandations de la Society for Psychophysiological Research note que le RMSSD sur des fenêtres de 10 secondes seulement et le SDNN sur des fenêtres de 30 secondes ne peuvent approcher les valeurs issues d'enregistrements longs que si l'on fait la moyenne de plusieurs fenêtres prises dans une période de repos [212]. Le RMSSD reflète la variabilité à court terme, et le SDNN la variabilité globale, qui dépend de la durée de l'enregistrement.

Tableau L6.1. Fiche de mesure des intervalles RR (modèle pour la feuille de calcul). Les formules sont données pour la ligne indiquée et se recopient vers le bas.

A : N° du battement	B : Instant R (s)	C : RR (ms)	D : FC (bpm)	E : Différence RR (ms)
1 (ligne 2)	valeur lue	non calculé	non calculé	non calculé
2 (ligne 3)	valeur lue	=(B3-B2)*1000	=60000/C3	non calculé
3 (ligne 4)	valeur lue	=(B4-B3)*1000	=60000/C4	=C4-C3
...	...	...	...	...
21 (ligne 22)	valeur lue	=(B22-B21)*1000	=60000/C22	=C22-C21

L6.3 Partie C. Expérience d'échantillonnage : le repliement

Le but est de voir de vos propres yeux ce qui se passe lorsque la fréquence d'échantillonnage descend sous le double de la fréquence du signal.

1. Dans une nouvelle feuille, placez dans la colonne A les nombres 0, 1, 2, ..., 199.
2. Dans la colonne B, le temps pour un échantillonnage à 100 Hz : `=A2/100`. Dans la colonne C, une sinusoïde de 10 Hz : `=SIN(2*PI()*10*B2)`. Tracez le graphique de C en fonction de B pour les 50 premières lignes. Comptez les cycles en une seconde.
3. Dans les colonnes E et F, refaites le calcul pour un échantillonnage à 12,5 Hz : le temps `=A2*0,08` (ou `=A2*0.08`, selon le séparateur décimal), la valeur `=SIN(2*PI()*10*E2)`. Tracez le graphique pour les 25 premières lignes (deux secondes).
4. Comptez de nouveau les cycles par seconde. Le graphique montre une onde de 2,5 Hz, c'est-à-dire |10 - 12,5| Hz, qui n'existe pas dans le signal réel. Comparez avec la figure 6.2 du chapitre.
5. Reliez le résultat à l'EEG : à 160 Hz, fréquence d'échantillonnage de la base EEG Motor Movement/Imagery, la plus haute fréquence restituée correctement est 80 Hz. Qu'arriverait-il à une interférence de 50 Hz si le même EEG était sous-échantillonné à 60 Hz sans filtre anti-repliement ?

La variante Python (facultative) réalise la même expérience et calcule automatiquement la fréquence dominante :

```
import numpy as np

fe = 1000                       # frequence d'echantillonnage initiale (Hz)
t = np.arange(0, 2, 1 / fe)     # 2 secondes
x = np.sin(2 * np.pi * 10 * t) # sinusoide de 10 Hz

for k in (10, 40, 80):          # on garde un echantillon sur k
    y = x[::k]
    fe_nou = fe / k
    spectru = np.abs(np.fft.rfft(y))
    frecv = np.fft.rfftfreq(len(y), d=1 / fe_nou)
    print(f"fe = {fe_nou:6.1f} Hz -> pic {frecv[spectru.argmax()]:.1f} Hz")
```

À l'exécution, le programme affiche un pic à 10 Hz pour l'échantillonnage à 100 Hz et à 25 Hz, mais à 2,5 Hz pour l'échantillonnage à 12,5 Hz. Ceux qui souhaitent travailler sur l'ECG réel peuvent lire les 10 premières secondes de l'enregistrement 100 directement sur PhysioNet avec les fonctions `wfdb.rdrecord('100', pn_dir='mitdb', sampto=3600)` et `wfdb.rdann('100', 'atr', pn_dir='mitdb', sampto=3600)`, documentées pour la version 4.3.1 du paquet [555]. Le sous-échantillonnage brut de l'ECG (`x[::k]`) déforme visiblement les complexes QRS.

L6.4 Partie D. EEG : le rythme alpha

1. Dans le dossier S001 de la base EEG Motor Movement/Imagery, ouvrez les fichiers `S001R01.edf` (yeux ouverts) et `S001R02.edf` (yeux fermés), d'environ 1,2 Mo chacun [553].
2. Ouvrez `S001R02.edf` dans EDFbrowser. Choisissez d'afficher les canaux occipitaux (O1, Oz, O2) et, à titre de comparaison, quelques canaux frontaux.
3. Réglez la fenêtre d'affichage sur 10 secondes. Recherchez sur les canaux occipitaux une activité rythmique, sinusoïdale, plus ample que sur les canaux frontaux.
4. Mesurez la fréquence : comptez les ondes dans un intervalle d'une seconde, à trois endroits différents du tracé, et faites la moyenne. Le rythme alpha postérieur est de 8-12 Hz ; il est le plus net les yeux fermés et s'atténue à l'ouverture des yeux [209].
5. Utilisez la fonction de spectre de puissance (FFT) d'EDFbrowser sur le canal O1, sur 30 à 60 secondes de signal [216]. Notez la fréquence du pic principal entre 8 et 12 Hz.
6. Refaites les étapes 2 à 5 pour `S001R01.edf`, les yeux ouverts. Comparez l'amplitude visuelle et la hauteur du pic spectral.
7. Observez aussi les artefacts : le clignement et les mouvements oculaires apparaissent surtout sur les canaux frontaux, et la contraction musculaire ajoute une activité irrégulière de haute fréquence. Certains artefacts physiologiques, comme les mouvements oculaires et l'activité musculaire, peuvent même aider à l'interprétation [219].

Le standard minimal IFCN-ILAE pour l'EEG clinique de routine recommande un échantillonnage d'au moins 256 Hz et un affichage avec des filtres de 0,5-70 Hz [208]. Le jeu de données utilisé ici, échantillonné à 160 Hz, est un jeu de recherche, adapté à l'exercice mais pas au diagnostic. Pour le rythme alpha, 160 Hz suffisent : la fréquence maximale restituée correctement est de 80 Hz, soit plus de six fois la limite supérieure de la bande alpha.

Pièges fréquents

- La mesure des intervalles RR en secondes et l'application de la formule FC = 60 000/RR, conçue pour des millisecondes.
- L'inclusion dans le calcul des intervalles qui entourent un battement A ou V. On obtient des valeurs de SDNN et de RMSSD fortement augmentées, qui ne décrivent plus la variabilité sinusale.
- La confusion entre le RR moyen et la moyenne des fréquences instantanées. Les deux ne donnent pas exactement la même fréquence cardiaque moyenne.
- La recherche du rythme alpha sur les canaux frontaux ou dans l'enregistrement les yeux ouverts.
- La lecture d'un pic spectral étroit à 50 ou 60 Hz (selon le pays où le signal a été enregistré) comme une activité cérébrale. Il s'agit de l'interférence du réseau électrique.

Tableau L6.2. Fiche d'observation du rythme alpha.

Enregistrement	Canal	Fréquence comptée (Hz)	Pic spectral (Hz)	Observations (amplitude, artefacts)
S001R02, yeux fermés	O1			
S001R02, yeux fermés	O2			
S001R01, yeux ouverts	O1			
S001R01, yeux ouverts	O2			

Travail à rendre

Chaque étudiant remet, au format électronique, un seul fichier de rapport (PDF ou document de la feuille de calcul), contenant :

1. Le tableau L6.1 rempli pour le segment propre et pour le segment contenant le battement A, avec le RR moyen, la FC moyenne, le SDNN et le RMSSD (3 points).
2. Le tachogramme du segment propre, avec des axes légendés et les unités (1 point).
3. Les graphiques de l'expérience de repliement et la fréquence apparente calculée (2 points).
4. Le tableau L6.2 rempli et une capture du spectre de puissance du canal O1 dans les deux conditions (2 points).
5. Des réponses courtes (5 lignes au maximum chacune) à trois questions : (a) pourquoi exclut-on le battement A du calcul de la VFC, (b) quelle fréquence d'échantillonnage minimale exigerait un signal comportant des composantes jusqu'à 150 Hz, (c) pourquoi le rythme alpha diminue-t-il à l'ouverture des yeux (1 point).
6. La citation correcte des bases de données et des logiciels utilisés (1 point).

Barème indicatif : 10 points. À l'examen pratique, l'étudiant peut recevoir un autre segment de l'enregistrement 100 ou un autre enregistrement de la base MIT-BIH, et il doit refaire les étapes de la partie L6.2 en 20 minutes.

Variante hors ligne

Si la salle n'a pas accès à Internet, l'enseignant télécharge à l'avance sur PhysioNet les fichiers `100.hea`, `100.dat` et `100.atr` de l'enregistrement 100 (environ 1,9 Mo au total) [213] ainsi que les fichiers `S001R01.edf` et `S001R02.edf` [553]. Les fichiers sont copiés sur les ordinateurs de la salle avec EDFbrowser. L'enregistrement ECG est converti en EDF+ à l'aide de l'utilitaire de conversion du format MIT intégré au logiciel, et les intervalles RR sont obtenus avec le détecteur de QRS intégré [216]. La partie C ne nécessite aucune connexion. Si EDFbrowser ne peut pas non plus être installé, la partie B se fait sur une série synthétique d'intervalles RR préparée par l'enseignant, par exemple celle de la figure L6.2, générée par le script `fig_L06.py`.

Travaux pratiques 7. Traitement des images médicales avec ImageJ/Fiji

Objectif

L'étudiant ouvre une image TDM au format DICOM, lit l'en-tête, applique une fenêtre de visualisation, effectue des mesures calibrées (longueur, aire, densité moyenne), compare deux filtres de réduction du bruit, segmente des objets par seuillage, les compte et exporte les résultats au format CSV. Toutes les étapes sont enregistrées dans une macro qui peut être exécutée de nouveau sur une autre image. La séance met en application les notions du chapitre 7 (sections 7.1, 7.4 et 7.6).

Outils

- **Fiji**, la distribution d'ImageJ qui inclut d'emblée de nombreux modules d'extension (*plugins*) pour l'analyse d'images [556]. Elle se télécharge gratuitement sur https://imagej.net/software/fiji/downloads (consulté le 27 septembre 2026). La variante recommandée est "Latest", qui inclut Java 21 et fonctionne sous Windows 10 ou version ultérieure, macOS 11 ou version ultérieure et Ubuntu 22.04 ou version ultérieure. La variante "Stable", avec Java 8, reste disponible pour les ordinateurs anciens [557].
- **Noyau ImageJ** : à la date de consultation, la dernière version figurant dans les notes de version officielles était la 1.54t, du 16 mai 2026 [558]. La version installée se vérifie dans "Help > About ImageJ".
- **Images** : les images de démonstration d'ImageJ, ouvertes depuis le menu "File > Open Samples", en particulier "CT (16-bit DICOM)" et "Blobs" [559]. On peut aussi utiliser une série publique de The Cancer Imaging Archive (TCIA), dont les collections sont anonymisées et, pour la plupart, disponibles sous licence Creative Commons Attribution 3.0 ou 4.0 pour un usage pédagogique [560]. On n'utilise pas d'images de patients réels de l'hôpital.
- **Tableur** (Excel, LibreOffice Calc ou Google Sheets) pour le fichier CSV exporté.

Déroulement

La figure L7.1 résume les huit étapes de la séance et les menus utilisés. Le tableau L7.1 peut rester ouvert à côté de Fiji pendant toute la séance.

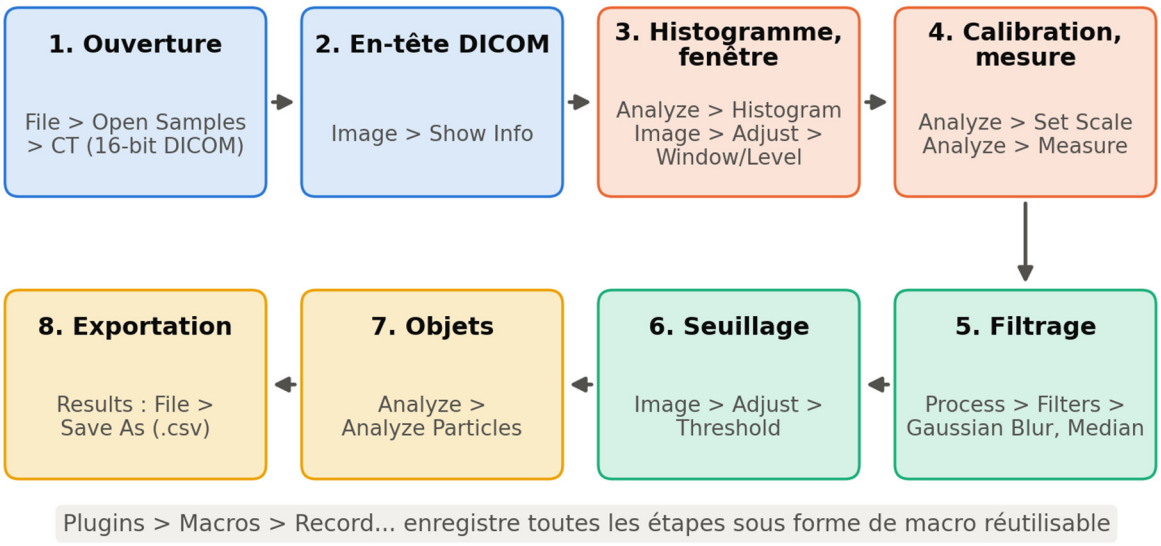


Figure L7.1. Déroulement du TP 7 dans Fiji, de l'ouverture de l'image DICOM à l'exportation des résultats. L'enregistreur de macros capture toutes les étapes.

Tableau L7.1. Opérations utilisées pendant la séance et leur chemin dans les menus de Fiji. Source : [556,559].

Opération	Chemin dans le menu	Raccourci	Ce que l'on note
Ouverture de l'exemple	"File > Open Samples > CT (16-bit DICOM)"	-	les dimensions de l'image
En-tête DICOM	"Image > Show Info"	Ctrl+I	les étiquettes (0028,0030), (0028,1050), (0028,1052)
Histogramme	"Analyze > Histogram"	Ctrl+H	la moyenne, le min, le max
Fenêtre	"Image > Adjust > Window/Level"	-	les valeurs de L et W utilisées
Calibration	"Analyze > Set Scale"	-	mm par pixel
Grandeurs mesurées	"Analyze > Set Measurements"	-	les colonnes qui apparaissent
Mesure	"Analyze > Measure"	Ctrl+M	longueur, aire, moyenne
Filtres	"Process > Filters > Gaussian Blur / Median"	-	sigma ou rayon
Seuil	"Image > Adjust > Threshold"	Ctrl+Shift+T	la méthode, les limites
Comptage des objets	"Analyze > Analyze Particles"	-	le nombre, l'aire moyenne
Macro	"Plugins > Macros > Record"	-	le fichier .ijm enregistré

L7.1 Installation et préparation

1. Téléchargez depuis la page officielle l'archive Fiji "Latest" adaptée à votre système d'exploitation [557].
2. Décompressez-la dans un dossier où vous disposez des droits d'écriture (par exemple `C:\LP7\Fiji` ou dans votre dossier personnel). Fiji ne nécessite pas d'installation classique.
3. Lancez l'exécutable Fiji depuis le dossier. Au premier démarrage, le programme peut proposer des mises à jour. Ne les acceptez que si l'assistant vous le demande, afin que tout le groupe travaille sur la même version.
4. Notez la version indiquée dans "Help > About ImageJ".
5. Créez le dossier de travail `C:\LP7\rezultate`, dans lequel vous enregistrerez tous les fichiers.

L7.2 Ouverture de l'image et lecture de l'en-tête DICOM

1. Ouvrez "File > Open Samples > CT (16-bit DICOM)". L'image est une coupe TDM téléchargée depuis le serveur d'ImageJ ; une connexion Internet est donc nécessaire.
2. Passez la souris sur l'image et observez la barre d'état de la fenêtre principale : les coordonnées et la valeur du pixel s'y affichent.
3. Ouvrez "Image > Show Info" (Ctrl+I). La fenêtre affiche l'en-tête DICOM sous la forme d'une liste d'étiquettes du type (groupe,élément), avec le nom de l'attribut et sa valeur.
4. Recherchez et recopiez sur la fiche de travail les valeurs de : Modality (0008,0060), Rows (0028,0010), Columns (0028,0011), Pixel Spacing (0028,0030), Bits Allocated (0028,0100), Bits Stored (0028,0101), Window Center (0028,1050), Window Width (0028,1051), Rescale Intercept (0028,1052) et Rescale Slope (0028,1053) [245]. Si un attribut manque, notez "absent".
5. Vérifiez la calibration des valeurs : placez le curseur dans l'air, à l'extérieur du patient. Si l'image est affichée en unités Hounsfield, la valeur doit se situer autour de -1 000. Si vous voyez de grandes valeurs positives, appliquez manuellement la relation UH = Rescale Slope · valeur stockée + Rescale Intercept pour trois pixels et notez le calcul.
6. Recherchez dans l'en-tête les attributs qui permettraient d'identifier un patient, par exemple (0010,0010) Patient's Name et (0010,0020) Patient ID. Notez s'ils sont vides, anonymisés ou renseignés. Discutez avec l'assistant de ce qu'il faudrait effacer avant qu'une telle image soit envoyée à un collègue pour un travail de recherche.

L7.3 Histogramme et fenêtre de visualisation

1. Ouvrez "Analyze > Histogram" (Ctrl+H). Observez les deux grands groupes de valeurs : l'air (autour de -1 000) et les tissus mous (autour de 0-100). Notez la moyenne, le minimum et le maximum affichés sous le graphique.
2. Ouvrez "Image > Adjust > Window/Level". Cliquez sur le bouton "Set" et saisissez le niveau 40 et la largeur 400 (fenêtre médiastinale). Décrivez ce que l'on voit au niveau des poumons.
3. Recommencez avec le niveau -600 et la largeur 1500 (fenêtre pulmonaire). Décrivez ce qu'il advient des tissus mous.
4. Calculez, avec la formule du chapitre 7, le niveau de gris d'un pixel de 20 UH dans chacune des deux fenêtres et comparez les résultats avec ce que vous observez à l'écran.
5. **N'appuyez pas sur le bouton "Apply".** Il modifie les valeurs des pixels, et pas seulement l'affichage. Si vous avez appuyé dessus par erreur, fermez l'image sans l'enregistrer et rouvrez-la.

L7.4 Calibration spatiale et mesures

1. Vérifiez la taille du pixel dans "Image > Properties". Pour un fichier DICOM, elle est reprise de Pixel Spacing, et l'unité doit être le mm.
2. Ouvrez "Analyze > Set Scale" uniquement pour voir la valeur. Ne la modifiez pas. Expliquez sur la fiche ce qu'il adviendrait des mesures si vous saisissiez 1 pixel = 1 mm.
3. Dans "Analyze > Set Measurements", cochez : "Area", "Mean gray value", "Standard deviation", "Min & max gray value" et "Perimeter".

4. Avec l'outil "Straight line", tracez le diamètre transversal du corps vertébral ou de l'aorte, selon le cas. Appuyez sur Ctrl+M. La longueur en mm apparaît dans la fenêtre "Results".
5. Avec l'outil "Oval", dessinez une région d'intérêt (ROI) à l'intérieur de l'aorte ou du foie, loin des bords. Appuyez sur Ctrl+M. Notez l'aire (mm²), la moyenne et l'écart type, en UH.
6. Ajoutez la ROI dans "Analyze > Tools > ROI Manager" (bouton "Add"), afin de pouvoir la réutiliser à l'identique à l'étape suivante.
7. Recommencez la mesure de la moyenne dans une ROI placée dans l'air. La valeur attendue est d'environ -1 000 UH [24].

L7.5 Filtrage du bruit

1. Dupliquez l'image ("Image > Duplicate") deux fois et nommez les copies "gauss" et "median".
2. Sur la copie "gauss", appliquez "Process > Filters > Gaussian Blur" avec sigma = 2 pixels.
3. Sur la copie "median", appliquez "Process > Filters > Median" avec un rayon de 2 pixels.
4. Appliquez sur chaque image la même ROI du ROI Manager et mesurez la moyenne et l'écart type.
5. Remplissez un tableau de trois lignes (original, gauss, median). L'écart type dans une région homogène donne une estimation du bruit, et le filtrage est l'une des opérations globales utilisées pour le réduire [24]. Le filtre gaussien lisse tout, y compris les bords, alors que le filtre médian élimine surtout les points isolés et préserve mieux les contours. Regardez le bord de l'aorte sur les deux copies et notez la différence.

L7.6 Segmentation par seuillage et comptage des objets

Le seuillage transforme une image en niveaux de gris en un masque binaire, objet contre fond. C'est une forme simple de segmentation, fréquemment utilisée, par exemple, pour écarter les structures qui recouvrent la zone d'intérêt lors du rendu volumique [24]. La figure L7.2 illustre le principe sur une image synthétique comportant 14 objets sur un fond bruité.

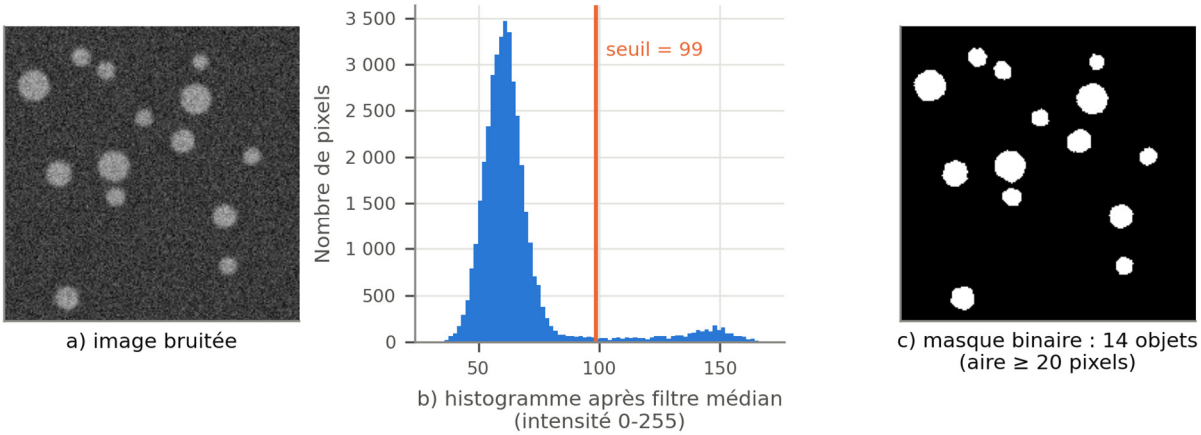


Figure L7.2. Segmentation par seuillage sur une image synthétique : a) l'image bruitée, b) l'histogramme après filtre médian, avec le seuil choisi automatiquement par la méthode d'Otsu, c) le masque binaire, dans lequel on compte 14 objets d'une aire d'au moins 20 pixels. Données synthétiques, illustratives.

1. Ouvrez "File > Open Samples > Blobs".
2. Ouvrez "Image > Adjust > Threshold". Choisissez la méthode "Default", puis essayez "Otsu". Observez la zone colorée en rouge, qui indique les pixels sélectionnés. Cochez ou décochez

l'option "Dark background" jusqu'à ce que ce soient les objets, et non le fond, qui soient sélectionnés.

3. Cliquez sur "Apply". L'image devient binaire (noir et blanc).
4. Ouvrez "Analyze > Analyze Particles". Réglez "Size (pixel^2)" sur 50-Infinity, cochez "Display results", "Summarize" et "Exclude on edges", puis cliquez sur "OK".
5. Relevez dans la fenêtre "Summary" le nombre d'objets ("Count") et l'aire moyenne ("Average Size"). Recommencez avec "Exclude on edges" décoché et expliquez la différence.
6. Facultatif, sur l'image TDM : avec "Image > Adjust > Threshold" et le bouton "Set", sélectionnez les valeurs de -1000 à -400 UH. Les pixels marqués correspondent à l'air et au poumon. Mesurez l'aire des poumons sur cette coupe avec "Analyze Particles". Comparez le résultat avec celui d'un collègue qui a choisi d'autres limites.

L7.7 Enregistrement d'une macro et exportation des résultats

Fiji peut être automatisé au moyen de macros, de scripts et, pour ceux qui programment, depuis Python [561,562]. L'enregistreur de macros transforme chaque commande de menu en une ligne de code.

1. Fermez toutes les images et les fenêtres de résultats. Lancez "Plugins > Macros > Record".
2. Refaites sur Blobs les étapes de L7.6 : filtre médian de rayon 2, seuil, "Apply", "Analyze Particles".
3. Dans la fenêtre "Recorder" apparaîtront des lignes semblables à celles ci-dessous (le code généré peut différer légèrement selon les options choisies) :

```
run("Median...", "radius=2"); setAutoThreshold("Default"); run("Convert to Mask");
run("Analyze Particles...", "size=50-Infinity display exclude summarize");
```

4. Cliquez sur "Create", enregistrez la macro sous `C:\LP7\rezultate\lp7_numarare.ijm` et exécutez-la sur une image nouvellement ouverte ("Macros > Run" ou Ctrl+R dans l'éditeur). Le résultat doit être identique à celui obtenu manuellement.
5. Dans la fenêtre "Results", choisissez "File > Save As" et enregistrez `C:\LP7\rezultate\lp7_blobs.csv`. Enregistrez de la même façon les mesures faites sur l'image TDM, sous le nom `lp7_ct.csv`.
6. Ouvrez les fichiers CSV dans le tableur. Calculez l'aire moyenne et l'écart type de l'aire des objets, puis comparez-les avec les valeurs de la fenêtre "Summary".

L7.8 Discussion : limites et reproductibilité

À la fin de la séance, le groupe compare les résultats. En général, l'aire des poumons obtenue à l'étape facultative de L7.6 diffère d'un étudiant à l'autre, bien que l'image soit la même. La différence ne vient pas du logiciel, mais des choix effectués : les limites du seuil, la forme et la position de la ROI, le filtre appliqué au préalable. Le même problème se pose, à une autre échelle, en radiomique, où les caractéristiques extraites des images varient avec l'appareil, le protocole et le logiciel utilisé [265]. C'est pourquoi l'initiative IBSI a défini et vérifié des valeurs de référence pour 169 caractéristiques, afin que des logiciels différents puissent être étalonnés les uns par rapport aux autres [266].

La macro enregistrée en L7.7 est la réponse pratique à cette variabilité : elle documente exactement ce qui a été fait et permet à quiconque de refaire l'analyse. Le fichier `.ijm` se remet avec les résultats, de même qu'un article scientifique décrit sa méthode.

Dans cette séance, Fiji est utilisé uniquement à des fins pédagogiques. Un logiciel utilisé pour poser un diagnostic doit satisfaire aux exigences applicables aux dispositifs médicaux [284], et, à l'hôpital, les images sont interprétées sur des stations de travail et des moniteurs étalonnés, connectés au PACS. Exporter une image en JPEG "pour la mesurer plus commodément" est une erreur fréquente : on perd la calibration spatiale et les valeurs UH, et donc la possibilité d'une mesure correcte.

Travail à rendre

Les fichiers sont déposés sur la plateforme de la discipline dans une archive `LP7_Nume_Prenume.zip`. Barème indicatif, sur 10 points :

- la fiche de travail avec les attributs DICOM de L7.2 et la vérification de la valeur de l'air (2 points) ;
- la description des deux fenêtres et le calcul du niveau de gris pour 20 UH (1,5 point) ;
- `lp7_ct.csv` avec la longueur, l'aire et la moyenne en UH, avec les unités correctes (2 points) ;
- le tableau original, gauss, median et le commentaire sur le bruit et les bords (1,5 point) ;
- `lp7_blobs.csv`, le nombre d'objets et l'explication de l'option "Exclude on edges" (1,5 point) ;
- la macro `lp7_numarare.ijm`, qui s'exécute sans erreur (1,5 point).

À l'examen pratique, l'étudiant peut recevoir une image inconnue ; il doit alors lire Pixel Spacing dans l'en-tête, mesurer une structure en mm et compter les objets par seuillage, en justifiant chaque choix.

Variante hors ligne

Si le laboratoire n'a pas accès à Internet, l'assistant copie avant la séance sur chaque ordinateur (ou sur une clé USB) l'archive Fiji téléchargée depuis la page officielle et les fichiers d'exemple `ct.dcm` et `blobs.gif`, que l'on ouvre ensuite avec "File > Open". Si les images d'exemple ne sont pas disponibles non plus, on génère une image synthétique équivalente à celle de la figure L7.2 en exécutant, dans "Plugins > New > Macro", le code suivant :

```
newImage("sintetic", "8-bit black", 256, 256, 1); setColor(150); fillOval(30, 40, 22, 22); fillOval(120, 60, 16, 16); fillOval(190, 150, 26, 26); fillOval(70, 180, 18, 18); run("Add Specified Noise...", "standard=18");
```

On applique à cette image les étapes de L7.5, L7.6 et L7.7. Le fond étant noir, on coche "Dark background" dans la boîte de dialogue "Threshold", et, dans la macro, la ligne de seuillage devient `setAutoThreshold("Default dark");`. Le nombre d'objets doit être de 4. La partie DICOM (L7.2-L7.4) se fait dans ce cas sur n'importe quel fichier DICOM anonymisé fourni par l'assistant, et les questions sur l'en-tête restent les mêmes.

Travaux pratiques 8. Travailler avec un dossier de santé électronique : démo OpenEMR et ressources FHIR

Objectif

Cette séance relie les deux faces du dossier de santé électronique (DSE) décrites au chapitre 8. La première est celle que voit le clinicien : des écrans, des formulaires, des listes. La seconde est celle que voient les systèmes informatiques : des ressources FHIR échangées au moyen de requêtes web. À la fin de la séance, l'étudiant sera capable :

- de créer dans OpenEMR un patient fictif et d'enregistrer une consultation avec des signes vitaux, un problème, une allergie et une prescription ;
- de formuler dans le navigateur des requêtes de recherche et de lecture adressées à un serveur FHIR public de test ;
- de reconnaître dans une ressource JSON les éléments principaux des ressources Patient et Observation (identifiant, nom, date, code LOINC, valeur, unité) ;
- de mettre en correspondance les écrans d'OpenEMR et les ressources FHIR ;
- de comparer, sur le plan conceptuel, le circuit d'une prescription dans OpenEMR et FHIR avec celui de la prescription électronique en Roumanie.

Règle de base de la séance : on n'utilise que des données fictives. Les serveurs employés sont publics, les données saisies sont visibles par les autres utilisateurs, et la saisie de données réelles d'un patient enfreindrait le secret professionnel et la législation relative à la protection des données.

Outils

- **Un navigateur web** à jour (Firefox, Chrome ou Edge). Firefox affiche par défaut les fichiers JSON dans une visionneuse structurée, ce qui facilite la partie B.
- **OpenEMR, démonstration en ligne officielle**, version 8.4.1, publiée le 20 septembre 2026 [563]. La page https://www.open-emr.org/demo/ (consultée le 27 septembre 2026) propose trois instances identiques et des comptes publics pour plusieurs rôles (administrateur, médecin, clinicien, comptable, accueil). Les mots de passe sont affichés sur la page et ne sont pas recopiés dans la fiche de travail. Chaque instance est remise à son état initial tous les jours à 8 h 00 UTC ; les données ne sont donc pas conservées [564]. OpenEMR est un logiciel open source et propose une API FHIR R4 conforme aux profils US Core, avec une autorisation SMART on FHIR [565].
- **Le serveur public de test HAPI FHIR**, point d'accès R4 : https://hapi.fhir.org/baseR4 (consulté le 27 septembre 2026, date à laquelle le logiciel du serveur indiquait la version de développement HAPI FHIR 8.13.10-SNAPSHOT) [566]. La page du serveur avertit qu'il ne

s'agit pas d'un serveur de production, qu'il ne faut pas y stocker d'informations de santé personnelles et que la base est vidée puis rechargée périodiquement avec des données de test [566].

- **La spécification FHIR**, pour le profil des signes vitaux et les codes LOINC correspondants [567].
- **Un éditeur de texte** (Notepad, TextEdit ou équivalent) pour la fiche de travail et pour copier les fragments JSON.

Déroulement

La séance comporte trois parties (figure L8.1). La durée indicative est de 30 minutes pour la partie A, de 35 minutes pour la partie B et de 25 minutes pour la partie C.

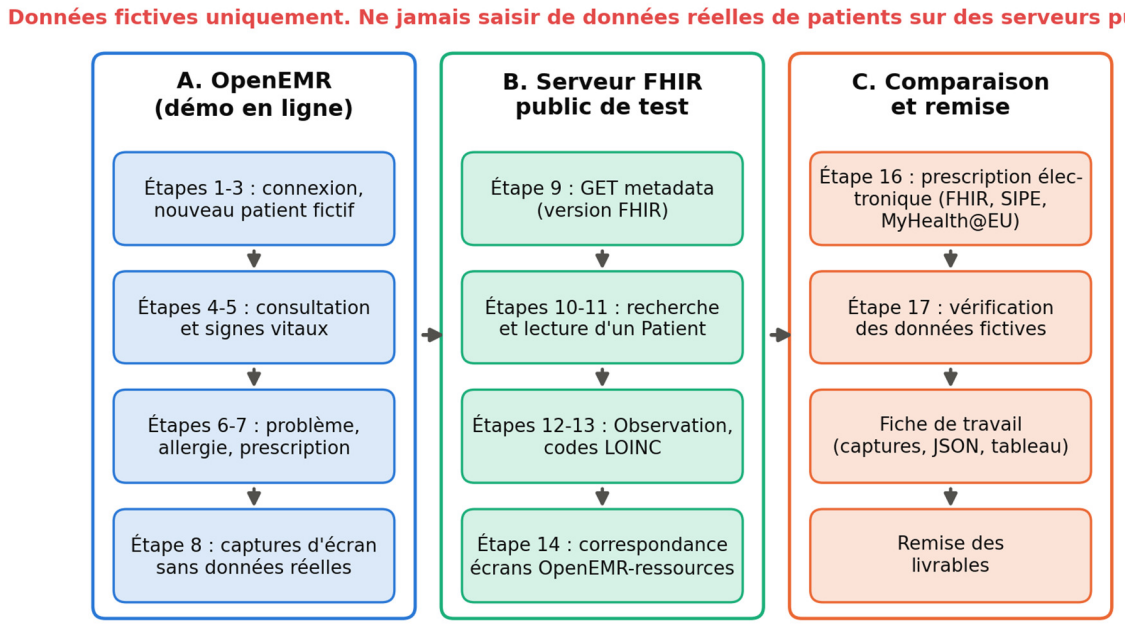


Figure L8.1. Déroulement des travaux pratiques : enregistrement d'un patient fictif dans OpenEMR (A), lecture des mêmes types de données sous forme de ressources sur un serveur FHIR public de test (B), comparaison avec la prescription électronique en Roumanie et remise du travail (C).

Chaque étudiant choisit un code personnel de travail, formé des initiales du groupe et d'un nombre (par exemple G3-17). Ce code entre dans le nom du patient fictif, afin que l'enregistrement puisse être retrouvé parmi ceux des collègues.

L8.1 Partie A : le patient fictif dans OpenEMR

1. Ouvrez https://www.open-emr.org/demo/, choisissez l'une des trois instances et connectez-vous avec le compte public de médecin ou d'administrateur indiqué sur la page. Si une instance est lente, passez à une autre.
2. Créez le patient. Dans le menu principal, choisissez *Patient* puis *New/Search*, remplissez les champs obligatoires et validez avec *Create New Patient* [568]. Utilisez les données fictives suivantes : nom de famille Testescu, prénom Ana-[votre code], date de naissance 12.05.1990, sexe féminin. Ne remplissez ni CNP (code numérique personnel roumain), ni téléphone, ni adresse réelle.

3. Observez le dossier du patient (*Dashboard*). Notez dans la fiche de travail les sections que vous voyez. La documentation d'OpenEMR décrit, entre autres, les sections *Medical Problems*, *Allergies*, *Medications* et *Prescriptions*, chacune dotée d'une icône de modification [568]. Les intitulés exacts peuvent légèrement varier d'une version à l'autre. Si vous ne trouvez pas un menu, cherchez-le d'après sa fonction, non d'après son nom.
4. Ouvrez une nouvelle consultation (*encounter*, appelée visite dans l'interface). Indiquez la date du jour, la catégorie de la visite et le motif : "Céphalées depuis 3 jours". Enregistrez.
5. Dans la consultation ouverte, ajoutez le formulaire des signes vitaux (*Vitals*) avec les valeurs suivantes : pression artérielle 150/95 mmHg, pouls 88/min, température 36,8 °C, poids 72 kg, taille 165 cm. Enregistrez et vérifiez si le logiciel a calculé l'indice de masse corporelle. Notez la valeur affichée.
6. À partir du dossier du patient, ajoutez un problème médical : "Hypertension artérielle essentielle", avec la date du jour comme date de début. Si l'interface propose la recherche d'un code de diagnostic, notez le code proposé. Ajoutez ensuite une allergie : pénicilline, réaction de type urticaire, sévérité modérée.
7. Ajoutez une prescription dans la section *Prescriptions* (en général par *Edit* puis *Add*) : amlodipine 5 mg, un comprimé par jour, 30 comprimés. Enregistrez.
8. Faites quatre captures d'écran : le dossier du patient avec le problème, l'allergie et la prescription visibles, le formulaire des signes vitaux, la liste des consultations et l'écran de la prescription. Vérifiez qu'aucune donnée personnelle réelle, la vôtre ou celle d'autres personnes, n'apparaît sur les captures.

Ce qu'il faut observer. Certaines données ont été saisies dans des champs structurés (les valeurs des signes vitaux, la date de naissance, éventuellement le code de diagnostic), d'autres en texte libre (le motif de la consultation). Notez dans la fiche de travail deux exemples de chaque catégorie. C'est la distinction entre documentation structurée et texte libre examinée à la section 8.1.2.

L8.2 Partie B : les mêmes concepts sous forme de ressources FHIR

La partie B n'utilise pas les données saisies dans OpenEMR. L'API FHIR d'OpenEMR exige une autorisation selon le standard SMART, que nous ne configurons pas au laboratoire [565]. Nous travaillons plutôt sur le serveur public HAPI FHIR, qui contient déjà des patients et des observations de test. Les requêtes s'écrivent directement dans la barre d'adresse du navigateur. Chaque requête GET est une lecture : elle ne modifie rien sur le serveur.

9. Vérifiez le serveur. Ouvrez https://hapi.fhir.org/baseR4/metadata?_format=json. La réponse est la déclaration de capacités (*CapabilityStatement*). Cherchez-y le champ `fhirVersion` et notez sa valeur (pour ce point d'accès, 4.0.1, c'est-à-dire R4).
10. Recherchez des patients d'après leur nom de famille : https://hapi.fhir.org/baseR4/Patient?family=smith&_count=5&_format=json. La réponse est une ressource *Bundle* de type `searchset`. Chaque élément de `entry` contient une ressource Patient. Le serveur renvoie les résultats par pages : le paramètre `_count` fixe le nombre de résultats par page, et l'élément `link` de relation `next`, s'il apparaît, contient l'adresse de la page suivante. Ouvrez cette adresse et observez qu'il s'agit, là encore, d'une requête GET. Choisissez ensuite un patient et notez la valeur de `id`.

11. Lisez le patient choisi : https://hapi.fhir.org/baseR4/Patient/[id]?_format=json, en remplaçant [id] par la valeur notée. Repérez dans le JSON les éléments `name` (avec `family` et `given`), `gender`, `birthDate` et, s'il existe, `identifier`. Copiez le fragment dans la fiche de travail.
12. Recherchez les observations de ce même patient : https://hapi.fhir.org/baseR4/Observation?patient=[id]&_count=10&_format=json. Si le résultat est vide, recherchez directement des observations de pouls d'après le code LOINC : https://hapi.fhir.org/baseR4/Observation?code=http://loinc.org|8867-4&_count=5&_format=json. Pour une observation, repérez :

- `code.coding` (le système, par exemple http://loinc.org, le code et le libellé) ;
- `valueQuantity` (la valeur, l'unité et le code de l'unité) ;
- `subject` (la référence au patient, de la forme Patient/[id]) ;
- `effectiveDateTime` (le moment de la mesure), s'il est présent ;
- `category` (par exemple `vital-signs` ou `laboratory`).

13. Reliez l'observation aux signes vitaux saisis dans la partie A. Dans le profil FHIR des signes vitaux, le pouls a le code LOINC 8867-4, le poids 29463-7, la température 8310-5, l'indice de masse corporelle 39156-5, tandis que la pression artérielle est un panel (85354-9) à deux composantes, la systolique (8480-6) et la diastolique (8462-4) [567]. Expliquez en deux phrases pourquoi la pression artérielle n'est pas représentée par une valeur unique.
14. Complétez le tableau L8.1, qui met en correspondance les écrans de la partie A et les ressources FHIR. Les colonnes consacrées aux éléments sont déjà remplies. Pour chaque ligne, notez dans la fiche de travail la valeur concrète saisie dans OpenEMR qui aboutirait dans l'élément correspondant.

Tableau L8.1. Correspondance entre les écrans d'OpenEMR et les ressources FHIR R4. Source : [565,567].

Écran OpenEMR	Ressource FHIR	Éléments principaux	Exemple tiré de la partie A
Données démographiques	Patient	name, gender, birthDate, identifier	Testescu, Ana-[code], 1990-05-12
Consultation (visite)	Encounter	status, period, reasonCode, subject	le motif "Céphalées depuis 3 jours"
Signes vitaux	Observation (vital-signs)	code (LOINC), valueQuantity, effectiveDateTime	pouls 88/min, code 8867-4
Problèmes médicaux	Condition	code, clinicalStatus, onsetDateTime, subject	hypertension artérielle essentielle
Allergies	AllergyIntolerance	code, reaction, patient	pénicilline, urticaire
Prescriptions	MedicationRequest	medication[x], dosageInstruction, authoredOn, requester	amlodipine 5 mg, 1 cp/jour

Toutes ces ressources figurent parmi celles qu'expose l'API FHIR d'OpenEMR [565]. Dans la version R4 utilisée ici, les éléments de liaison s'appellent `subject` (pour Observation, Condition, Encounter, MedicationRequest) et `patient` (pour AllergyIntolerance). Cette différence de dénomination est l'un des pièges habituels lorsqu'on écrit des recherches.

15. Facultatif, pour les étudiants qui terminent plus tôt : ajoutez à la requête de l'étape 12 le paramètre `&_summary=count`. La réponse ne contient que le nombre total de résultats. Expliquez à quoi servirait une telle requête dans une application qui affiche le nombre d'analyses d'un patient.

Un avertissement pratique. Sur le serveur HAPI, n'importe quel utilisateur peut écrire, et les données peuvent être modifiées ou supprimées à tout moment. Ne supposez pas qu'un `id` donné existera encore la semaine prochaine et ne tirez aucune conclusion clinique de ces données. N'utilisez au laboratoire aucune requête POST, PUT ou DELETE : elles écriraient sur un serveur public.

L8.3 Partie C : comparaison avec la prescription électronique en Roumanie

Dans la partie A, la prescription a été enregistrée dans un seul logiciel. En FHIR, cette même prescription deviendrait une ressource MedicationRequest, et sa délivrance en pharmacie une ressource distincte, MedicationDispense, que l'API d'OpenEMR expose également [565]. La norme ISO 17523 fixe les données que doit contenir une prescription électronique, mais laisse à chaque pays la transmission, la validation et le remboursement [312].

En Roumanie, la prescription électronique circule par l'intermédiaire du Sistemul informatic de prescripție electronică (système informatique de prescription électronique, SIPE), qui fait partie de la plateforme de la CNAS (Caisse nationale d'assurance maladie), et, depuis le 1er septembre 2026, le patient peut consulter ses ordonnances des 5 dernières années sur le portail eSănătateaMea [322,324]. Au niveau européen, les prescriptions et les délivrances électroniques sont deux des catégories prioritaires de l'espace européen des données de santé, avec un échange obligatoire entre les États membres à partir du 26 mars 2029 [23].

16. Répondez par écrit, en 150 à 200 mots, aux questions suivantes :

- Quels sont les acteurs du circuit (médecin, pharmacien, patient, caisse d'assurance maladie) et quelle information chacun reçoit-il ?
- Quelle ressource FHIR correspondrait à l'ordonnance émise et quelle ressource correspondrait à la délivrance en pharmacie ?
- Que faudrait-il, sur les plans technique et organisationnel, pour que l'ordonnance émise à Timișoara puisse être délivrée dans une pharmacie d'un autre État membre ? Utilisez les niveaux d'interopérabilité de la figure 8.2.

17. Vérifiez encore une fois la fiche de travail et les captures. Aucune information ne doit permettre d'identifier une personne réelle.

Travail à rendre

On remet un seul document (PDF ou DOCX), avec le nom et le groupe de l'étudiant sur la première page. Le barème est indicatif, sur 10 points :

- 1 point d'office ;
- 2 points : les quatre captures d'écran de la partie A, avec le patient fictif, la consultation, les signes vitaux, le problème, l'allergie et la prescription visibles ;
- 1 point : les deux exemples de données structurées et les deux exemples de texte libre (étape 8) ;
- 2 points : les adresses exactes des requêtes FHIR des étapes 9 à 12 et les fragments JSON copiés, avec les éléments demandés signalés (soulignés ou surlignés) ;
- 1 point : l'explication de la représentation de la pression artérielle (étape 13) ;
- 1 point : le tableau L8.1 complété avec vos propres valeurs (étape 14) ;
- 2 points : la réponse à l'étape 16, correcte et argumentée.

À l'examen pratique, l'étudiant peut recevoir un fragment JSON inconnu (une ressource Patient ou Observation) et doit indiquer, en quelques minutes, le patient, la date, le code, la valeur et l'unité.

Variante hors ligne

Si le laboratoire n'a pas accès à Internet, la partie A est remplacée par un exercice sur papier et la partie B par des fichiers locaux.

1. L'enseignant enregistre à l'avance, sur les ordinateurs du laboratoire, deux exemples officiels de la spécification FHIR R4 : https://hl7.org/fhir/R4/patient-example.json et https://hl7.org/fhir/R4/observation-example-heart-rate.json (adresses vérifiées le 27 septembre 2026). Les fichiers s'ouvrent avec un éditeur de texte ou avec le navigateur, sans connexion. En secours, on peut utiliser les exemples Patient et HL7 v2 du chapitre 8.
2. Les étudiants remplissent sur papier un formulaire de consultation pour le patient fictif de la partie A, avec les mêmes champs (données démographiques, motif, signes vitaux, problème, allergie, prescription), et signalent les champs structurés et ceux en texte libre.
3. Sur les fichiers locaux, on parcourt les étapes 11 à 13 : repérage des éléments dans la ressource Patient et dans la ressource Observation du pouls, puis rédaction d'un fragment JSON de type Observation pour la température de 36,8 °C du patient fictif, avec le code LOINC 8310-5 et l'unité Cel [567].
4. Le tableau L8.1 et la réponse à l'étape 16 sont complétés de la même manière. Le barème est conservé, avec le formulaire sur papier à la place des captures d'écran et le fragment JSON écrit à la main à la place des requêtes au serveur.

Si le laboratoire dispose d'une installation locale d'OpenEMR préparée par l'enseignant, la partie A se fait sur celle-ci, avec les mêmes données fictives.

Travaux pratiques 9. Simulation d'une téléconsultation : préparation, consentement, documentation et limites

Objectif

Les étudiants travaillent en binômes et simulent une téléconsultation vidéo complète, dans laquelle l'un joue le rôle du médecin et l'autre celui d'un patient fictif. À la fin de la séance, chaque étudiant doit être capable de préparer une téléconsultation à l'aide d'une liste de vérification, d'obtenir et de consigner le consentement, de reconnaître les signes d'alarme qui imposent l'arrêt de la consultation à distance, de rédiger la note de téléconsultation selon le modèle de la section 9.6 et d'évaluer de façon critique la note du collègue.

La séance n'a pas pour but l'apprentissage d'une technique clinique. Son but est la discipline informationnelle de l'acte médical à distance : qui est le patient, où il se trouve, par quel canal nous communiquons, ce que nous ne pouvons pas voir et comment la trace écrite est conservée. Les études menées auprès d'étudiants montrent que ces savoir-faire n'apparaissent pas d'eux-mêmes. Dans une comparaison portant sur 241 étudiants de deuxième année, ceux qui avaient été évalués lors d'une consultation simulée par télémédecine ont obtenu des scores plus faibles pour l'examen physique et la communication interpersonnelle que ceux évalués en face à face, mais des scores plus élevés pour l'anamnèse [569]. Un module de deux heures consacré à la communication par vidéo a accru la confiance déclarée des étudiants de dernière année [570].

L'enjeu est clinique, et pas seulement formel. Dans une cohorte de 2 393 patients de la Mayo Clinic, le diagnostic provisoire posé par vidéo a été confirmé lors de l'examen en présentiel dans 86,9 % des cas [342]. L'écart d'environ 13 % explique pourquoi la note doit indiquer clairement ce qui n'a pas pu être examiné et quand le médecin reverra le patient. La séance fait pratiquer précisément ces deux gestes.

Aucune donnée réelle de patient n'est utilisée. Tous les cas sont fictifs et signalés comme tels, et les consultations ne sont pas enregistrées.

Outils

- **Une plateforme de visioconférence approuvée par la faculté** : on utilise le compte institutionnel dont les étudiants disposent déjà. Comme solution de remplacement gratuite, à code source ouvert, on peut utiliser Jitsi Meet (version stable 2.0.11248, publiée le 14 septembre 2026, https://meet.jit.si, consulté le 27 septembre 2026) [571]. Sur le serveur public meet.jit.si, la personne qui crée la salle doit avoir un compte, alors que les participants n'en ont pas besoin [340]. C'est pourquoi les salles sont créées par l'enseignant, qui envoie les liens aux binômes.

- **Les paramètres de sécurité** : l'étudiant vérifie dans la documentation de la plateforme si le flux est chiffré uniquement en transit ou également de bout en bout. Dans Jitsi Meet, le chiffrement de bout en bout ne fonctionne que dans certains navigateurs et ne couvre pas les messages du chat [340].
- **Les documents** : le modèle de note de téléconsultation (section 9.6 du chapitre 9), la liste de vérification préalable à la visite (tableau L9.1), la fiche de cas fictif remise par l'enseignant, la grille d'évaluation par les pairs.
- **Un éditeur de texte** : Google Docs ou Microsoft Word, dans le compte institutionnel. La note est remise au format PDF.
- **La référence de travail** : le guide de l'OMS pour la mise en œuvre de la télémédecine, pour l'identification, le consentement et les procédures standard [334].

Déroulement

L9.1 La préparation (15 minutes)

1. Les binômes se forment au début de la séance. Chaque étudiant jouera tour à tour les deux rôles, avec deux cas différents.
2. L'enseignant distribue les fiches de cas. La fiche du "patient" indique l'âge, le motif, les symptômes, les antécédents et, pour certains cas, un signe d'alarme qui n'est révélé qu'en réponse à une question appropriée. La fiche du "médecin" ne contient que le motif du rendez-vous. Ce modèle s'inspire des scénarios de patient standardisé utilisés pour entraîner la conscience de la situation (*situational awareness*) en télémédecine [572].

Les cas de travail (tous fictifs) sont choisis de sorte que chaque binôme rencontre au moins une situation qui ne peut pas être résolue à distance :

- **Cas A** : femme de 58 ans, hypertendue, contrôle programmé. Les valeurs mesurées à domicile au cours de la dernière semaine dépassent la cible, sans symptômes. Le cas peut être résolu à distance, avec un ajustement du traitement et une réévaluation.
- **Cas B** : homme de 34 ans présentant une lésion pigmentée apparue récemment dans le dos, qu'il ne parvient pas à bien montrer à la caméra. La bonne décision est l'examen en présentiel ou une téléexpertise avec une image dermatoscopique obtenue dans un cabinet.
- **Cas C** : homme de 62 ans qui demande une "ordonnance pour des brûlures d'estomac". Interrogé sur l'effort, il rapporte une douleur rétrosternale avec sueurs, apparue il y a une heure. La consultation à distance s'arrête et le patient est immédiatement orienté vers le 112.
- **Cas D** : étudiante de 21 ans, traitée pour une dépression, visite de suivi. L'évaluation peut se faire par vidéo, mais la question sur les idées suicidaires est obligatoire et doit être consignée.

3. Le "médecin" parcourt la liste de vérification du tableau L9.1 avant d'entrer dans la salle : l'identité, la localisation, le numéro de téléphone de secours, le consentement, la confidentialité du lieu, le plan en cas de coupure de la connexion.
4. Le "patient" choisit un endroit calme, éclairé de face, et garde à portée de main un "tensiomètre" imaginaire dont il lit les valeurs sur la fiche.

Tableau L9.1. Liste de vérification avant et au début de la téléconsultation. Source : élaboré d'après [334,370].

Moment	Élément vérifié	Comment le vérifier	À consigner
Avant	la plateforme et la caméra fonctionnent	test audio-vidéo 5 minutes avant	non
Début	identité du patient	nom complet et date de naissance, comparés à la fiche	oui
Début	identité du médecin	nom, spécialité, établissement	oui
Début	localisation du patient	adresse au moment de la consultation	oui
Début	numéro de téléphone de secours	le patient le dicte, le médecin le répète	oui
Début	confidentialité	qui d'autre est dans la pièce, le patient peut-il parler librement	oui
Début	consentement	explication des limites et accord verbal	oui
Début	enregistrement	on annonce que la consultation n'est pas enregistrée	oui
En cours	signes d'alarme	questions ciblées au début de l'anamnèse	oui
Fin	plan et sécurité	recommandations, signes devant amener à reconsulter, délai	oui

L9.2 Le consentement et le cadre de la consultation (5 minutes)

5. Le "médecin" se présente, confirme l'identité du "patient" et sa localisation.
6. Il explique, en termes simples, ce qui peut et ce qui ne peut pas se faire à distance. Exemple de formulation : "Je peux vous écouter, vous voir et lire les valeurs que vous avez mesurées. Je ne peux pas ausculter votre cœur ni vos poumons. Si je le juge nécessaire, je vous demanderai de venir au cabinet ou d'appeler le 112."
7. Il demande l'accord verbal pour poursuivre la consultation à distance. Le guide de l'OMS précise qu'une consultation initiée par le professionnel exige un consentement documenté et que les limites de la télémédecine doivent être communiquées explicitement [334]. En Roumanie, les normes relatives à la télémédecine exigent, hors urgence, le consentement exprès du patient, prouvé par tout moyen de preuve [364]. L'accord verbal est donc consigné dans la note, avec l'heure à laquelle il a été obtenu.
8. Il fixe le plan de secours : "Si la communication est coupée, je vous rappelle dans les deux minutes au numéro que vous m'avez donné."

L9.3 La consultation proprement dite (10-12 minutes)

9. **Le triage initial.** Les premières questions recherchent les signes d'alarme : douleur thoracique, dyspnée de repos, faiblesse brutale d'un membre, troubles de la parole, saignement important, altération de l'état de conscience, idées suicidaires. Si l'un d'eux apparaît, la consultation à distance s'arrête. Le "médecin" indique d'appeler le 112 ou de se présenter immédiatement aux urgences et reste en ligne jusqu'à la prise en charge, si la situation l'exige (figure L9.1).
10. **L'anamnèse** suit la structure habituelle : le motif, l'histoire de la maladie actuelle, les antécédents, les médicaments, les allergies.
11. **L'examen guidé par vidéo.** Le "médecin" demande au "patient" des gestes simples : s'approcher de la caméra, montrer ses chevilles, compter à voix haute pendant 30 secondes, lire les valeurs de sa pression artérielle. Les revues consacrées à l'examen physique virtuel décrivent justement ce type de manœuvres, avec leurs limites [339].

12. **La décision.** Le "médecin" juge si le problème peut être résolu à distance. Si ce n'est pas le cas, il programme un examen en présentiel en précisant la date et le lieu.
13. **La conclusion.** Le "médecin" résume le plan, vérifie la compréhension (il demande au "patient" de répéter ce qu'il doit faire) et formule les conseils de sécurité : quels signes imposent de reconsulter ou d'appeler le 112. Ces conseils sont souvent omis de la documentation [371].
14. En cours de séance, l'enseignant peut "couper" la connexion d'un binôme (en demandant au "patient" d'éteindre sa caméra). Le "médecin" doit appliquer le plan de secours annoncé.

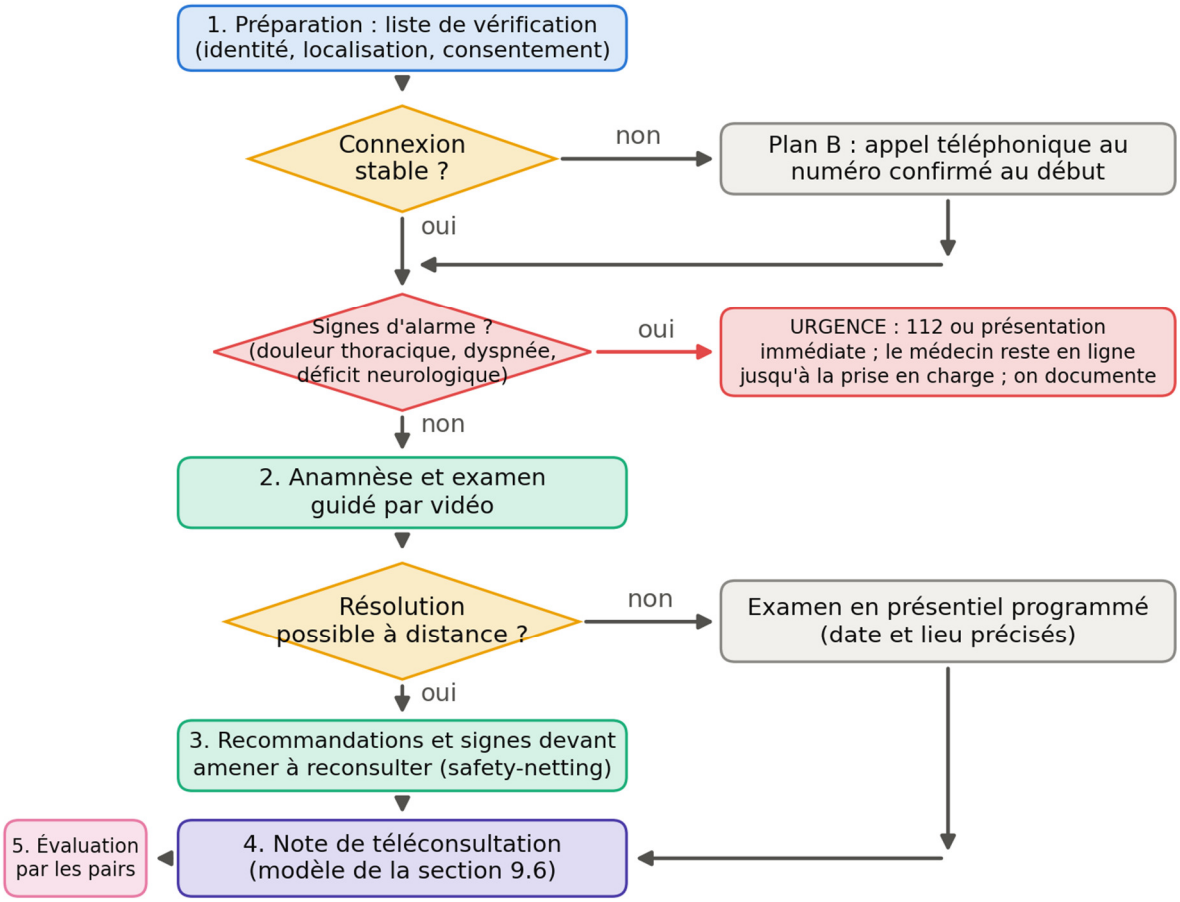


Figure L9.1. Déroulement de la téléconsultation simulée, avec les points de décision : stabilité de la connexion, signes d'alarme et possibilité d'une résolution à distance. Schéma didactique élaboré d'après [334].

L9.4 La rédaction de la note (15 minutes)

15. Immédiatement après la consultation, le "médecin" rédige la note de téléconsultation selon le modèle de la section 9.6 : identification, localisation, modalité, consentement, données cliniques, limites, évaluation et plan, sécurité, signature. Si la simulation a comporté une coupure de la connexion, l'incident technique est consigné, comme l'exigent également les normes roumaines [364]. On écrit ce qui a effectivement été observé, et non ce qu'aurait montré un examen complet.
16. La note ne contient aucune donnée réelle. Les noms sont des initiales fictives, et le CNP (code numérique personnel) est remplacé par la mention "[fictif]".
17. Les rôles s'inversent et les étapes 5 à 16 sont répétées avec le second cas.

L9.5 L'évaluation par les pairs (15 minutes)

18. Les binômes échangent leurs notes. Chaque étudiant évalue la note de son collègue à l'aide de la grille ci-dessous, en cochant chaque élément présent et correct.
19. Grille d'évaluation par les pairs (un point par élément) :

- l'identité du patient et du médecin, la date et l'heure de début et de fin.
- la localisation du patient et le numéro de téléphone de secours.
- la modalité, la plateforme et les personnes présentes.
- le consentement à la consultation à distance et la mention relative à l'enregistrement.
- les données cliniques, clairement séparées entre ce que rapporte le patient et ce qu'observe le médecin.
- les limites de l'examen, formulées concrètement.
- l'évaluation et le plan, y compris le délai de réévaluation.
- les signes d'alarme expliqués au patient.
- la bonne décision de triage pour le cas reçu (poursuite à distance, examen en présentiel ou urgence).
- l'absence de données d'identification réelles et d'informations inutiles.

20. L'évaluateur rédige deux observations : un élément bien réalisé et un élément à corriger. Les outils élaborés pour évaluer les compétences en télémédecine en simulation utilisent eux aussi des grilles par compétence, remplies par des enseignants ou par des patients standardisés, même si leur validation reste préliminaire [573].
21. Discussion finale avec tout le groupe (5 minutes) : quels signes d'alarme ont été manqués, quelles limites ont été omises dans les notes, qu'est-ce qui a mal fonctionné sur le plan technique.

L9.6 Le lien avec le TP 10

Les notes anonymisées produites au cours de cette séance sont conservées. Lors du TP 10, les étudiants demanderont à un grand modèle de langage (*large language model*, LLM) de résumer ou de structurer une note de téléconsultation fictive et vérifieront, avec la même grille, quels éléments le modèle a omis ou inventés. L'exercice prépare à une question qu'ils rencontreront dans la pratique : qui répond d'une note rédigée à l'aide d'un outil automatique ?

Les compétences exercées ici correspondent aux domaines des référentiels internationaux de compétences numériques pour l'enseignement médical, qui incluent le professionnalisme en santé numérique et les systèmes d'information de santé [30], ainsi qu'aux compétences en télémédecine proposées par consensus Delphi pour la formation postuniversitaire, dont la moitié concernent la pratique fondée sur le système et le professionnalisme [574].

Travail à rendre

Chaque étudiant remet, au format PDF, dans un seul fichier :

1. la liste de vérification remplie pour la consultation dans laquelle il a joué le rôle du médecin (tableau L9.1).
2. la note de téléconsultation rédigée selon le modèle de la section 9.6.
3. la grille d'évaluation remplie pour la note du collègue, avec les deux observations.
4. un paragraphe de 80 à 120 mots sur une limite de la consultation à distance observée pendant la simulation.

Tableau L9.2. Barème indicatif pour l'examen pratique (10 points).

Critère	Points
Liste de vérification complète et correcte	2
Note de téléconsultation : identification, localisation, modalité, consentement	2
Note : données cliniques, limites, évaluation et plan	2
Note : conseils de sécurité et décision de triage correcte	2
Évaluation par les pairs argumentée	1
Paragraphe de réflexion	1

Une décision de triage erronée dans le cas comportant un signe d'alarme (poursuite de la consultation à distance) entraîne la perte des 2 points du quatrième critère, quelle que soit la qualité du reste de la note.

Variante hors ligne

Si la plateforme vidéo ou le réseau ne fonctionnent pas, la simulation se fait en face à face, les étudiants étant assis dos à dos. Le "médecin" n'a pas le droit de se retourner. On simule ainsi une téléconsultation téléphonique, dans laquelle l'image manque elle aussi. L'examen guidé se réduit à ce que le patient peut rapporter, et l'étudiant doit consigner dans la note cette limite supplémentaire. La liste de vérification, la note et la grille sont remplies sur papier, sur les formulaires imprimés par l'enseignant, puis photographiées ou numérisées pour être remises. La discussion finale compare les deux situations : ce qui a été perdu lorsque l'image a disparu et quelles questions l'ont compensé.

Travaux pratiques 10. La consultation simulée avec un modèle de langage : vérification, erreurs et déclaration

Objectif

L'étudiant utilise un grand modèle de langage (*large language model*, LLM) comme interlocuteur dans une consultation simulée, sur un cas fictif, et traite chaque réponse comme une hypothèse à vérifier. À la fin, chaque étudiant doit être capable de rédiger une instruction structurée (*prompt*) pour un cas clinique, de décomposer la réponse en affirmations vérifiables, de vérifier chaque affirmation dans un guide de pratique clinique ou dans une source primaire, de classer les erreurs selon une grille publiée, de vérifier dans PubMed les références proposées par le modèle et de rédiger une déclaration d'utilisation de l'IA conforme aux recommandations de l'ICMJE.

Ce TP répond à un décalage documenté au chapitre 10. Testés seuls sur dix scénarios, trois modèles ont identifié une affection pertinente dans 94,9 % des cas en moyenne, mais les personnes qui les ont utilisés y sont parvenues dans moins de 34,5 % des cas, sans avantage par rapport au groupe témoin [403]. La différence est venue de la conversation : les utilisateurs donnaient souvent des informations incomplètes et ne suivaient pas toujours les suggestions correctes. La performance dépend aussi de la quantité et de l'ordre des informations reçues [402]. Un étudiant qui sait formuler le cas complet et vérifier la réponse est moins exposé à ces erreurs.

Ce TP prolonge le TP 9. Ceux qui ont conservé la note de téléconsultation fictive rédigée lors de ce TP peuvent l'utiliser comme matériel complémentaire à l'étape L10.4.

On travaille uniquement sur des cas fictifs, signalés comme tels. Aucune information concernant un patient réel n'est saisie dans le chatbot, pas même "seulement avec les initiales", car les données concernant la santé constituent une catégorie particulière protégée par l'article 9 du RGPD [118].

Outils

- **Un chatbot LLM accessible gratuitement dans le navigateur.** On privilégie le compte institutionnel, si la faculté en fournit un. Le TP ne dépend d'aucun produit particulier. L'étudiant note le nom de l'outil, la version du modèle affichée dans l'interface, la date et l'heure de la séance et si l'outil avait accès à la recherche sur Internet. Ce sont les données exigées par la déclaration de l'étape L10.7.
- **PubMed** (https://pubmed.ncbi.nlm.nih.gov, consulté le 27 septembre 2026), pour la vérification des références et la recherche des sources primaires [39].
- **Guides officiels gratuits** : les lignes directrices de l'Organisation mondiale de la santé approuvées par son comité d'examen des lignes directrices [575] et les recommandations des

sociétés savantes européennes, téléchargées au format PDF depuis les sites officiels. Les blogs, les résumés commerciaux et les autres réponses générées par l'IA ne servent pas de sources de vérification.

- **Les recommandations de l'ICMJE** (mises à jour en janvier 2026), sections consacrées aux outils d'IA [202].
- **La fiche de travail** : tableau comportant les colonnes "affirmation", "source de la vérification", "verdict", "catégorie de l'erreur", "gravité", remis par l'enseignant dans un format modifiable (Google Sheets ou Excel).

Déroulement

La durée totale est d'environ 110 minutes. Les étudiants travaillent individuellement, et la discussion finale se déroule avec l'ensemble du groupe.

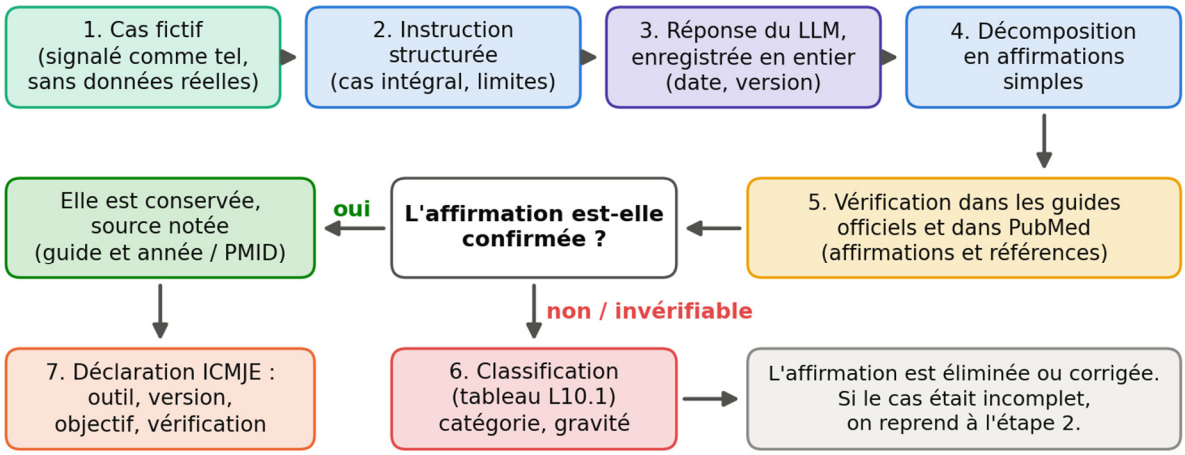


Figure L10.1. Boucle de vérification de la réponse d'un modèle de langage au cours du TP : du cas fictif et de l'instruction structurée, en passant par la décomposition en affirmations, la vérification dans les sources et la classification des erreurs, jusqu'à la déclaration d'utilisation. Toute affirmation non confirmée revient dans la boucle ou est éliminée. Schéma didactique élaboré d'après [202,406].

L10.1 Le choix du cas (5 minutes)

L'enseignant distribue l'un des cas fictifs ci-dessous. Chaque cas contient au moins un élément qui modifie la conduite à tenir et qui peut passer inaperçu.

- **Cas A (fictif).** Femme de 71 ans, hypertendue et diabétique, dont la montre connectée a affiché trois fois au cours de la dernière semaine une notification de rythme irrégulier. Elle a des palpitations intermittentes, sans syncope. Sa question : "Est-ce que je dois faire quelque chose ou c'est juste la montre ?"
- **Cas B (fictif).** Homme de 45 ans présentant une céphalée apparue brutalement pendant un effort, qu'il décrit comme "la plus forte de ma vie", accompagnée de nausées. Il demande quel antalgique prendre.
- **Cas C (fictif).** Étudiante de 20 ans souffrant d'un mal de gorge et de fièvre depuis deux jours, sans toux. Elle demande un antibiotique "pour que ça passe plus vite" avant la session d'examens.

L10.2 L'instruction structurée (15 minutes)

L'étudiant rédige une instruction en six parties : le rôle demandé au modèle, le contexte, le cas complet, la tâche, le format de la réponse et les limites. Des instructions vagues produisent des réponses vagues, et le modèle comble les informations manquantes par des suppositions. C'est pourquoi le cas est transcrit intégralement, avec l'âge, le sexe, la durée des symptômes, les comorbidités et le traitement en cours.

Exemple

- Rôle : "Réponds comme un médecin de famille qui explique à un étudiant de première année."
- Contexte : "Le cas est fictif et sert à un exercice didactique."
- Cas : le texte intégral du cas, sans aucune donnée personnelle réelle.
- Tâche : "Quels sont les diagnostics à envisager, quels signes d'alarme faut-il rechercher et quelle conduite recommandes-tu ?"
- Format : "Numérote les affirmations. Pour chaque affirmation, indique la source (guide, année) et, pour les articles, le PMID ou le DOI."
- Limites : "Dis explicitement ce que tu ne peux pas évaluer sans examen clinique et quand le patient doit consulter un médecin en urgence."

L'instruction est enregistrée exactement sous la forme envoyée. Lors de l'évaluation, on regarde si elle demandait explicitement les signes d'alarme et les limites.

L10.3 L'enregistrement de la réponse (10 minutes)

1. On ouvre une nouvelle conversation, sans historique, et on envoie l'instruction une seule fois.
2. La réponse est copiée intégralement dans la fiche de travail, avec la date, l'heure, le nom de l'outil et la version du modèle.
3. On envoie ensuite une deuxième question, contenant une prémisse fausse introduite volontairement par l'enseignant (par exemple, une affirmation erronée sur l'équivalence de deux médicaments). L'étudiant note si le modèle corrige la prémisse ou l'accepte. Dans une étude portant sur cinq modèles de pointe, l'acceptation de demandes illogiques de ce type a atteint 100 % [407].
4. On ne régénère pas la réponse pour en obtenir une "meilleure". La variabilité d'une exécution à l'autre est discutée à la fin, à partir des réponses différentes obtenues par les camarades pour le même cas.

L10.4 La décomposition en affirmations et la vérification (35 minutes)

1. La réponse est découpée en affirmations simples, une par ligne dans la fiche de travail. Exemple : "la notification de la montre confirme la fibrillation atriale" est une affirmation. "Un ECG est recommandé" en est une autre.
2. Pour chaque affirmation, l'étudiant cherche une confirmation dans un guide officiel ou dans un article trouvé dans PubMed. Dans la fiche, on inscrit le titre du guide, l'année, la page ou la section, ou bien le PMID de l'article.
3. Le verdict prend trois valeurs : confirmée, infirmée, invérifiable. Une affirmation invérifiable n'est pas considérée comme correcte.
4. On compare ensuite la réponse au cas. Toutes les données ont-elles été reprises (âge, comorbidités, durée des symptômes) ? Manque-t-il un élément qui modifierait la conduite à

tenir ? Les omissions sont inscrites dans la fiche sur des lignes distinctes, par exemple "n'a pas posé de question sur le traitement anticoagulant".

5. Ceux qui ont apporté la note du TP 9 peuvent demander au modèle de la résumer et vérifier, phrase par phrase, si le résumé contient un élément absent de la note ou s'il a omis un point important. Cette procédure reproduit, à petite échelle, la méthode utilisée dans une étude d'évaluation des notes cliniques générées automatiquement, dans laquelle des cliniciens ont vérifié chaque phrase de la note par rapport à la transcription de la consultation [406].

L10.5 La classification des erreurs (15 minutes)

Chaque affirmation infirmée et chaque omission reçoit une catégorie et une gravité. La grille du tableau L10.1 part de la taxonomie proposée par Asgari et ses collaborateurs pour les notes cliniques générées par un LLM : hallucinations de type fabrication, négation, causalité ou contexte, et omissions, chacune étant classée comme majeure ou mineure selon qu'elle pourrait modifier ou non le diagnostic ou la conduite à tenir [406]. Dans cette étude, le taux d'hallucinations a été de 1,47 % des 12 999 phrases des notes, et celui des omissions de 3,45 % des phrases des transcriptions [406]. Pour la consultation, la grille ajoute trois catégories didactiques : l'information obsolète, le conseil à risque et le signe d'alarme manqué. Cette dernière catégorie repose sur des données empiriques : un outil de triage destiné au grand public a sous-estimé la gravité de 52 % des urgences de référence lors d'un test structuré [404].

Tableau L10.1. Grille de classification des erreurs, avec des exemples fictifs. Source : adapté de [404,406,408].

Catégorie	Définition de travail	Exemple fictif	Gravité typique
Omission	il manque une information qui modifie la conduite	ne pose pas de question sur le traitement anticoagulant	majeure ou mineure
Fabrication	affirmation ou source sans correspondant réel	cite une "étude de 2024" avec un PMID inexistant	majeure
Négation	affirme le contraire des données du cas	"la patiente n'a pas de facteurs de risque", alors qu'elle a une HTA et un diabète	majeure
Causalité	lien de causalité non étayé	attribue les palpitations au café sans argument	mineure ou majeure
Contexte	information correcte dans un autre contexte	applique la conduite de l'adulte à un cas pédiatrique	majeure
Information obsolète	recommandation issue d'une version remplacée	utilise un score de risque tiré d'un ancien guide	mineure ou majeure
Conseil à risque	recommandation susceptible de nuire	suggère un antibiotique sans indication	majeure
Signe d'alarme manqué	absence d'orientation vers les urgences	céphalée d'installation brutale renvoyée à une consultation dans quelques jours	majeure

Pour le cas B, une recommandation de prendre un antalgique et d'attendre, sans orientation vers les urgences, est classée comme signe d'alarme manqué, de gravité majeure, quelle que soit la justesse du reste de la réponse.

L10.6 La vérification des références (15 minutes)

1. Si le modèle n'a pas donné de références, on lui en demande cinq, avec PMID ou DOI, pour les affirmations principales.
2. Chaque référence est recherchée dans PubMed par son PMID et, séparément, par son titre [39].
3. La référence reçoit l'un des quatre verdicts suivants : correcte (elle existe et étaye l'affirmation), réelle mais avec des données erronées (auteurs, année, revue), réelle mais sans rapport avec l'affirmation, inexistante.
4. On calcule le pourcentage de références inexistantes dans sa propre réponse et on l'inscrit dans la fiche.

Les résultats sont ensuite comparés à la littérature. Dans une étude portant sur 636 références, 55 % de celles générées par GPT-3.5 et 18 % de celles générées par GPT-4 n'existaient pas, et une partie des références réelles comportaient des erreurs substantielles [408]. Dans un autre test, les taux de références hallucinées ont été de 39,6 % pour GPT-3.5, de 28,6 % pour GPT-4 et de 91,4 % pour Bard [44]. Un audit mené en 2026 sur cinq chatbots populaires a constaté qu'aucun n'avait produit une liste de références entièrement correcte [405]. Un pourcentage nul dans son propre exercice ne prouve pas que l'outil est sûr. Il montre seulement que, cette fois-ci, les références ont résisté à la vérification.

L10.7 La déclaration d'utilisation (10 minutes)

Les recommandations de l'ICMJE exigent que l'utilisation des outils d'IA soit déclarée. L'aide à la rédaction est mentionnée dans la section des remerciements, tandis que l'utilisation pour la collecte ou l'analyse des données ou pour les figures est décrite dans les méthodes, de manière suffisamment détaillée pour pouvoir être reproduite, avec l'outil, la version et les instructions données. Les chatbots ne peuvent pas être auteurs, les humains répondent de l'ensemble du contenu, et le contenu généré par l'IA n'est pas cité comme source primaire [202]. L'étudiant rédige une déclaration de 60 à 100 mots pour son rapport. Voici un modèle possible :

Exemple

- "Pour ce rapport, j'ai utilisé [nom de l'outil], version [version affichée], le [date], afin de générer une réponse à un cas clinique fictif, à partir de l'instruction figurant en annexe 1."
- "Toutes les affirmations et références de la réponse ont été vérifiées dans des guides officiels et dans PubMed. Les résultats de la vérification sont présentés en annexe 2."
- "Le texte du rapport a été rédigé par l'auteur, sans aide de l'IA à la rédaction. L'auteur assume l'entière responsabilité du contenu."

L10.8 La réflexion et la discussion finale (10 minutes)

Chaque étudiant rédige 100 à 150 mots sur trois questions. Quelle erreur aurait pu nuire à un patient réel ? Quelle information du cas a le plus modifié la réponse du modèle ? L'erreur aurait-elle été remarquée sans vérification ? La discussion en groupe compare les réponses différentes obtenues pour le même cas ainsi que le nombre de références inexistantes. L'enseignant relie les conclusions aux orientations de l'OMS sur les grands modèles multimodaux, qui mettent en garde contre les affirmations fausses ou incomplètes et contre le biais d'automatisation, c'est-à-dire la tendance de l'utilisateur à ne plus remarquer les erreurs du système [410].

Travail à rendre

Chaque étudiant rend un seul fichier PDF, qui contient :

1. l'instruction structurée, exactement sous la forme envoyée ;
2. la réponse intégrale du modèle, avec la date, l'heure, l'outil et la version ;
3. la fiche de vérification : les affirmations, la source de chaque vérification, le verdict, la catégorie et la gravité de l'erreur ;
4. le tableau de vérification des références, avec le pourcentage de références inexistantes ;
5. la déclaration d'utilisation de l'IA ;
6. la réflexion de 100 à 150 mots.

Tableau L10.2. Barème indicatif pour l'examen pratique (10 points).

Critère	Points
Instruction structurée complète (cas intégral, signes d'alarme, limites)	1,5
Décomposition correcte en affirmations	1,5
Vérification avec des sources officielles identifiables (guide et année ou PMID)	2
Classification des erreurs et de leur gravité selon la grille	2
Vérification des références dans PubMed	1,5
Déclaration d'utilisation conforme aux recommandations de l'ICMJE	1
Réflexion	0,5

La saisie dans le chatbot de toute donnée réelle concernant un patient entraîne l'annulation du travail. Une affirmation marquée "confirmée" sans source identifiable est notée comme non vérifiée.

Variante hors ligne

Si le réseau ou le chatbot ne sont pas disponibles, l'enseignant distribue sur papier trois réponses générées auparavant par un LLM pour les cas A, B et C, avec les instructions utilisées, la date et la version du modèle. Il distribue également des extraits imprimés des guides pertinents, ainsi que les notices PubMed des références citées par le modèle, mélangées à des références qui n'existent pas dans PubMed. Les étudiants parcourent les étapes L10.4-L10.8 sur ces documents. La vérification des références se fait en comparant la liste générée aux notices imprimées. Les titres et les PMID qui ne figurent pas dans le jeu de notices sont notés "à vérifier" et recherchés dans PubMed à la première occasion où l'on dispose d'un accès au réseau. La déclaration d'utilisation est rédigée pour un rapport hypothétique qui utiliserait la réponse obtenue.

Travaux pratiques 11. Sécurité pratique : mots de passe, authentification multifacteur, vérification d'une signature électronique et reconnaissance de l'hameçonnage

Objectif

À la fin du TP, l'étudiant sera capable de calculer l'entropie d'un mot de passe généré aléatoirement et de la comparer aux recommandations actuelles, d'activer la validation en deux étapes sur son propre compte Google, de signer et de vérifier un document PDF fictif avec un certificat autosigné, de lire le certificat TLS d'un site, de reconnaître les signaux d'alarme d'un courriel d'hameçonnage (*phishing*) et de décrire les étapes du signalement d'un incident. Le TP applique les notions du chapitre 11 et la formule de l'entropie du chapitre 2.

Outils

- Calculatrice de poche ou Google Sheets (la fonction `LOG(x; 2)` pour le logarithme en base 2).
- Un compte Google personnel de l'étudiant, qu'il administre lui-même. Page d'aide officielle sur la validation en deux étapes : https://support.google.com/accounts/answer/185839, consultée le 27 septembre 2026 [533].
- Adobe Acrobat Reader, gratuit, version du canal continu 26.002.21931 (mise à jour du 17 septembre 2026, d'après les notes de version disponibles à l'adresse https://www.adobe.com/devnet-docs/acrobatetk/tools/ReleaseNotesDC/index.html, consultées le 27 septembre 2026). Téléchargement : https://get.adobe.com/reader/. Pages d'aide utilisées : la création d'un identifiant numérique autosigné [576] et la validation des signatures [577].
- Google Chrome 154 (passé sur le canal stable le 22 septembre 2026, d'après https://chromereleases.googleblog.com, consulté le 27 septembre 2026) ou Mozilla Firefox, avec la page d'aide de Chrome sur les symboles de sécurité [578].
- Le navigateur des listes de confiance de la Commission européenne, https://eidas.ec.europa.eu/efda/trust-services/browse/eidas/tls, consulté le 27 septembre 2026 [579].

On ne fait aucune capture d'écran contenant l'adresse électronique, le numéro de téléphone ou les codes de secours de l'étudiant. Tous les documents et messages du TP sont fictifs.

Déroulement

Le TP comporte six parties, de 10 à 35 minutes chacune, qui se déroulent dans l'ordre. La partie L11.3 est celle qui fait l'objet de l'évaluation la plus détaillée.

L11.1 La robustesse d'un mot de passe (20 minutes)

Pour un mot de passe choisi réellement au hasard, chaque caractère est un événement à N issues équiprobables. Comme on l'a vu au chapitre 2, un tel événement apporte log2 N bits. Un mot de passe de L caractères a donc pour entropie :

H = L · log2 N (bits)

1. Calculez log2 N pour trois alphabets : lettres minuscules (N = 26, log2 26 ≈ 4,70), lettres majuscules, minuscules et chiffres (N = 62, log2 62 ≈ 5,95), l'ensemble des 94 caractères imprimables du clavier (log2 94 ≈ 6,55).
2. Notez dans votre cahier l'entropie pour : 8 lettres minuscules (8 · 4,70 ≈ 37,6 bits), 8 caractères alphanumériques (≈ 47,6 bits), 15 lettres minuscules (≈ 70,5 bits), 12 caractères parmi 94 (≈ 78,7 bits).
3. Une phrase de passe composée de mots choisis au hasard dans une liste de 7 776 mots apporte log2 7 776 ≈ 12,9 bits par mot. Calculez l'entropie pour 5 et pour 6 mots (≈ 64,6 et ≈ 77,5 bits).
4. Estimez le temps moyen nécessaire pour deviner le mot de passe par essais successifs, dans l'hypothèse didactique de 10^10 essais par seconde : t = 2^H / 2 / 10^10 secondes. Les résultats approximatifs sont : 10 secondes pour 8 lettres minuscules, environ 3 heures pour 8 caractères alphanumériques, environ 45 ans pour 5 mots et environ 2 700 ans pour 15 lettres minuscules.
5. Comparez avec le guide NIST SP 800-63B-4 : un mot de passe utilisé comme facteur unique compte au moins 15 caractères, et celui qui est utilisé avec un second facteur au moins 8. Le guide n'exige plus de combinaisons obligatoires de types de caractères ni de changement périodique, mais il exige que le mot de passe soit vérifié par rapport à une liste de valeurs compromises ou très utilisées [445].
6. Rédigez deux phrases de conclusion. Remarquez que la longueur augmente davantage l'entropie que la complexité et que la formule n'est valable que pour un choix aléatoire. Un mot de passe inventé par un humain ("Spital2026!") a une entropie réelle bien plus faible que ne le laisse penser le nombre de caractères, parce qu'il suit des schémas prévisibles.

L11.2 La validation en deux étapes sur le compte Google (15 minutes)

Les étapes suivent la page d'aide officielle de Google [533].

1. Ouvrez votre compte Google (myaccount.google.com) et sélectionnez la rubrique consacrée à la sécurité et à la connexion.
2. Dans la rubrique relative à la façon dont vous vous connectez à Google, choisissez d'activer la validation en deux étapes et suivez les instructions à l'écran.
3. Comme seconde étape, privilégiez les invites Google sur le téléphone ou une clé d'accès (*passkey*). La page d'aide précise que les invites contribuent à protéger le compte contre le détournement du numéro de téléphone (*SIM swap*) [533].
4. Générez les codes de secours et conservez-les hors ligne, en lieu sûr. Ne les photographiez pas et ne les envoyez à personne.
5. Si vous utilisez un compte de la faculté ou un compte professionnel, les étapes peuvent différer, et la page de Google recommande de contacter l'administrateur [533].

6. Notez dans votre cahier, sans données personnelles, les méthodes de validation proposées par le compte et celle que vous avez choisie. Rattachez ce choix à la recommandation du NIST selon laquelle les applications doivent proposer une option résistante à l'hameçonnage, tandis que le code envoyé par SMS reste un authentificateur restreint [445].

L11.3 Signature et vérification d'un PDF fictif dans Acrobat Reader (35 minutes)

Le TP utilise un certificat autosigné, créé par l'étudiant. Il suffit pour observer le mécanisme, et la page d'Adobe le recommande pour un usage personnel ou pour des processus internes, sans autorité de certification [576]. Il n'a toutefois pas la valeur d'une signature qualifiée, car il ne repose pas sur un certificat qualifié délivré par un prestataire de services de confiance [433].

1. Créez dans Google Docs ou Word une "lettre médicale" d'une page pour un patient fictif (par exemple "Patient Test, 70 ans"), avec un diagnostic, un traitement et la posologie d'un médicament. Exportez-la au format PDF.
2. Dans Acrobat Reader, créez l'identifiant numérique autosigné : Menu > Preferences > Signatures, puis More dans Identities & Trusted Certificates, Digital IDs, le bouton d'ajout, l'option "A new digital ID I want to create now" [576]. Choisissez l'algorithme RSA sur 2 048 bits, et non sur 1 024 (la norme FIPS 186-5 exige pour RSA un module d'au moins 2 048 bits) [427]. Protégez l'identifiant par un nouveau mot de passe.
3. Signez le PDF avec l'outil Use a certificate, option de signature numérique (l'intitulé exact peut varier d'une version à l'autre), et enregistrez une copie signée.
4. Ouvrez le panneau Signatures et lancez la validation (Options > Validate Signatures). Ouvrez Signature Properties, puis Show Signer's Certificate [577]. Notez : le nom du titulaire, l'émetteur (identique au titulaire, puisque le certificat est autosigné), la période de validité, l'algorithme et l'état affiché.
5. Observez l'état affiché. Pour confirmer l'authenticité du signataire, on vérifie si son certificat ou les certificats parents sont dignes de confiance [577]. Un certificat autosigné ne remonte pas à une autorité de certification ; c'est pourquoi, sur l'ordinateur d'une personne qui ne l'a pas déclaré digne de confiance, le document peut apparaître intact, mais l'identité du signataire reste non confirmée.
6. Ajoutez un commentaire ou une note sur le document signé, enregistrez et lancez à nouveau la validation. Notez ce qu'Acrobat signale à propos des modifications faites après la signature.
7. Vérifiez l'horodatage. Si la signature ne comporte pas d'horodatage provenant d'un serveur de confiance, l'heure provient de l'horloge de l'ordinateur du signataire [577].
8. Parcourez sur papier le schéma de la figure L11.1 pour votre signature et pour un scénario hypothétique : une lettre signée par un médecin au moyen d'un certificat qualifié, délivré par un prestataire figurant sur la liste de confiance d'un État de l'UE. Dans le navigateur des listes de confiance [579], recherchez la Roumanie et notez combien de prestataires de services de confiance qualifiés y figurent à la date du TP. Lien avec le texte du règlement : seule la signature qualifiée a un effet juridique équivalent à celui d'une signature manuscrite [433].

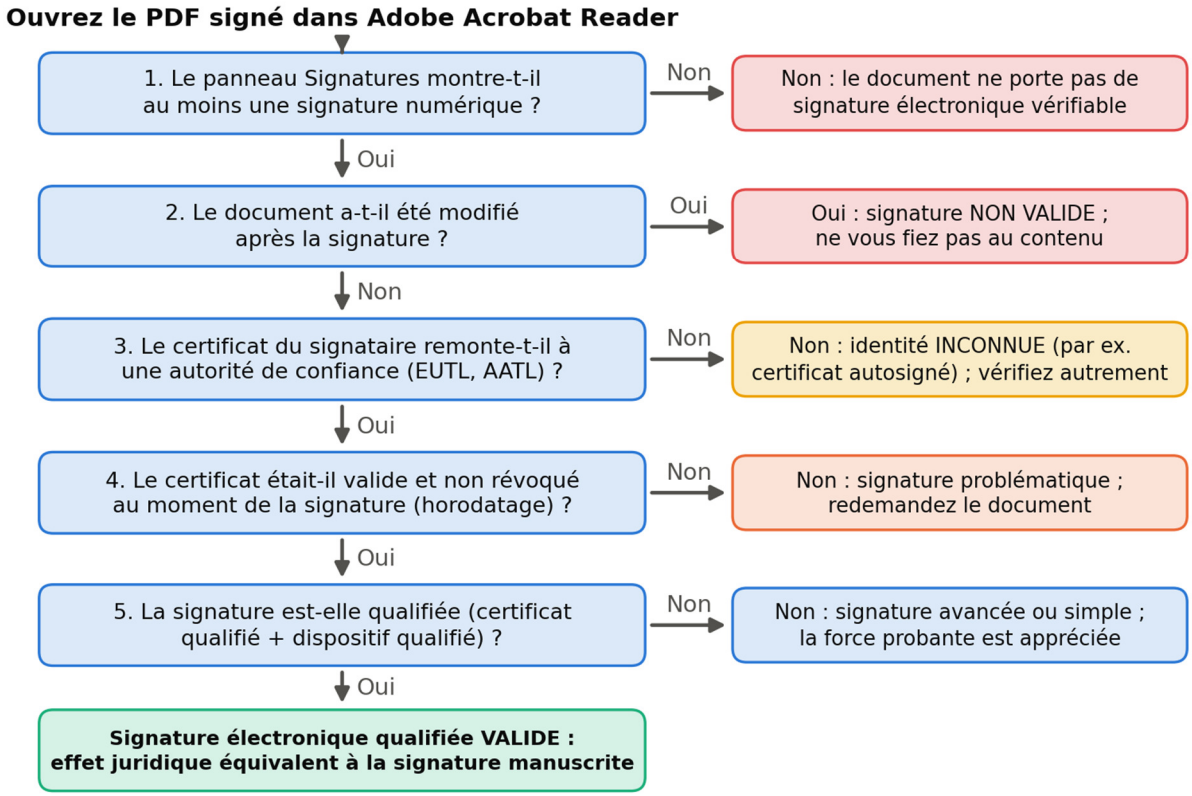


Figure L11.1. Schéma de décision pour la vérification d'une signature électronique sur un PDF dans Acrobat Reader. Adapté de [433,577].

L11.4 Le certificat TLS d'un site (10 minutes)

1. Ouvrez dans Chrome le site public d'une institution, par exemple https://cnas.ro.
2. Regardez le symbole situé à gauche de l'adresse. La page d'aide de Chrome explique la signification des symboles de connexion sécurisée, non sécurisée ou dangereuse [578].
3. Cliquez sur le symbole, puis sur les informations relatives à la connexion et sur le certificat. Notez : le nom pour lequel le certificat a été délivré, l'émetteur, la période de validité et la chaîne de certification jusqu'à l'autorité racine.
4. Comparez le nom figurant dans le certificat avec l'adresse de la barre du navigateur. Discutez en groupe des raisons pour lesquelles un cadenas affiché ne prouve pas que le site appartient à l'institution qu'il imite (un site d'hameçonnage peut lui aussi avoir un certificat valide).

L11.5 Reconnaître l'hameçonnage : test avec des courriels fictifs (25 minutes)

L'enseignant projette ou distribue les cinq messages ci-dessous. Tous sont fictifs, et les domaines se terminent par .example, un suffixe réservé qui n'appartient à personne. Les étudiants travaillent individuellement pendant 10 minutes et indiquent, pour chaque message, "hameçonnage" ou "légitime", en s'appuyant sur les signaux du tableau L11.1. La discussion suit.

1. De "Caisse d'assurance - Vérification" <alerta@cas-verificare.example>. Objet : "Votre carte de santé sera suspendue dans 24 heures". Le texte demande de confirmer ses données et son code PIN au moyen d'un lien raccourci.

2. De "Service informatique" <it-support@spital-judetean.example>, l'adresse interne habituelle. Objet : "Maintenance programmée samedi, 02:00-04:00". Il ne contient pas de lien et ne demande aucune donnée ; il annonce seulement l'indisponibilité du système.
3. De "Pr X, chef de service" <sef.clinica.x@mail-gratuit.example>. Objet : "Urgent et confidentiel". Il demande d'acheter cinq cartes-cadeaux et d'en envoyer les codes, sans téléphoner, parce que "je suis en réunion".
4. De "Ressources humaines" <hr@spita1-judetean.example> (le chiffre 1 à la place de la lettre l). Pièce jointe "Grila_salarii_2026.xlsm" qui demande l'activation des macros.
5. De la plateforme d'enseignement en ligne de la faculté, à l'adresse correcte. Message concernant un nouveau devoir. Le lien affiché et le lien réel (visible au survol avec la souris) coïncident et mènent au domaine de la faculté.

Réponses pour l'enseignant : 1, 3 et 4 relèvent de l'hameçonnage, 2 et 5 sont légitimes. Le message 3 illustre la fraude par usurpation de l'identité d'un supérieur, et le message 4 utilise un domaine d'apparence presque identique. Dans un hôpital italien de plus de 6 000 salariés, les messages personnalisés en fonction du contexte local ont été consultés bien plus souvent que les messages génériques (55 % contre 7 % des messages envoyés, lors de la première campagne) [443]. C'est pourquoi les messages 3 et 4 sont les plus dangereux.

Tableau L11.1. Signaux d'alarme dans un message d'hameçonnage. Source : synthèse didactique ; les messages du test sont fictifs.

Signal	Ce qu'il faut vérifier	Exemple dans le test
Urgence ou menace	délais très courts, sanctions, "dernier avertissement"	message 1
Demande de données secrètes	mots de passe, codes PIN, codes reçus par SMS	message 1
Expéditeur suspect	domaine gratuit ou différent du domaine officiel	message 3
Domaine contrefait	lettres remplacées (1 au lieu de l, rn au lieu de m)	message 4
Lien masqué	l'adresse réelle diffère du texte affiché, lien raccourci	message 1
Pièce jointe à risque	fichiers avec macros, archives, exécutables	message 4
Autorité et secret	un supérieur demande un paiement et interdit toute vérification	message 3
Contexte inhabituel	demande qui ne correspond pas aux procédures	messages 1, 3, 4

L11.6 La procédure de signalement d'un incident (10 minutes)

Par groupes de trois, les étudiants rédigent une procédure de 10 lignes au maximum pour le service dans lequel ils effectuent leur stage, à partir du scénario suivant : "Un interne a saisi son mot de passe sur la page du message 1". La procédure doit comprendre :

1. Ce qu'il ne faut pas faire : ne plus cliquer sur le lien, ne pas transférer le message aux collègues, ne pas supprimer le message avant de l'avoir signalé.
2. Ce qu'il faut faire immédiatement : changer le mot de passe depuis un autre appareil, prévenir le responsable informatique ou le responsable de la sécurité de l'hôpital selon la procédure interne, et déconnecter l'équipement suspect du réseau si le service informatique le demande.

3. Qui d'autre doit être prévenu : le délégué à la protection des données, si des données de patients ont pu être exposées (chapitre 5).
4. Au niveau national, les incidents de cybersécurité peuvent être signalés à la Direction nationale de la cybersécurité (Directoratul Național de Securitate Cibernetică, DNSC) par la plateforme PNRISC ou au numéro 1911, disponible 24 heures sur 24 [447].
5. Ce que l'on en retient : le guide de l'ENISA pour la santé recommande un processus de réponse aux incidents avec des rôles définis et la connaissance de la personne de contact, et, pour le personnel, des simulations d'hameçonnage et une formation adaptée aux rôles [446].

Travail à rendre

Chaque étudiant rend un fichier PDF nommé "L11_Nom_Prenom.pdf", qui contient :

1. Le tableau d'entropie de L11.1, avec les calculs et la conclusion (2 points).
2. La confirmation écrite, sans captures comportant des données personnelles, que la validation en deux étapes est active, avec la méthode choisie et sa justification (1 point).
3. Le PDF fictif signé, accompagné de la fiche d'observations de L11.3 : les données du certificat, l'état de la validation avant et après la modification, l'explication de la raison pour laquelle une signature avec certificat autosigné n'est pas une signature qualifiée (3 points).
4. Les données du certificat TLS de L11.4 (1 point).
5. Les réponses au test d'hameçonnage, avec au moins un signal du tableau L11.1 pour chaque message (2 points).
6. La procédure de signalement de L11.6 (1 point).

Total : 10 points. À l'examen pratique, l'étudiant peut être invité à calculer l'entropie d'un mot de passe donné ou à interpréter l'état d'une signature dans un PDF.

Variante hors ligne

En l'absence d'accès à Internet, les calculs de L11.1 et le test d'hameçonnage de L11.5 se font entièrement sur papier. Pour L11.3, Acrobat Reader installé au préalable permet de créer l'identifiant autosigné, de signer et de valider sans connexion. Pour L11.2 et L11.4, l'enseignant présente les étapes à l'aide du schéma de la figure L11.1 et d'un certificat TLS exporté au préalable sous forme de fichier, que les étudiants ouvrent et analysent en local. La procédure de L11.6 est rédigée sur papier.

Travaux pratiques 12. Le flux d'information au cabinet, en polyclinique et à l'hôpital : modélisation et indicateurs

Objectif

L'étudiant modélise le parcours d'un patient fictif, du médecin de famille jusqu'au remboursement des services, sous la forme d'un diagramme en couloirs (*swimlane*). Pour chaque étape, il identifie le document émis et les éléments de données minimaux, puis calcule dans un tableur quatre indicateurs hospitaliers sur un jeu de données fictif : la durée moyenne de séjour, le taux d'occupation des lits, l'indice de case-mix et le taux de réhospitalisation à 30 jours. Le TP applique les notions du chapitre 12 (sections 12.4 et 12.6).

Outils

- **draw.io** (appelé aussi diagrams.net), logiciel gratuit et libre (licence Apache 2.0), utilisable sans compte. La version web s'ouvre à l'adresse https://app.diagrams.net, et les diagrammes peuvent être enregistrés sur l'ordinateur personnel ou dans un espace de stockage choisi par l'utilisateur ; l'éditeur déclare ne pas avoir accès aux données enregistrées [580]. La version de bureau pour Windows, macOS et Linux se télécharge depuis la page officielle des versions. Au 27 septembre 2026, la dernière version était la 31.5.3, publiée le 26 septembre 2026 [581]. Les instructions officielles pour les diagrammes en couloirs figurent dans la documentation de draw.io [582].
- **Google Sheets** (dans le navigateur, avec un compte Google) ou **Microsoft Excel**. Les fonctions utilisées sont SUM, COUNT, COUNTA, AVERAGE, AVERAGEIF, COUNTIF, COUNTIFS, VLOOKUP et SUMPRODUCT. Leur description officielle se trouve dans la liste des fonctions de Google Sheets [583] et dans la documentation de Microsoft [584], toutes consultées le 27 septembre 2026.
- **Le jeu de données fictif** du tableau L12.2, recopié à la main ou importé depuis un fichier CSV fourni par l'enseignant. Aucune donnée de patient réel n'est utilisée.

Le travail avec les formules du tableur suit le protocole de [6], repris et mis à jour pour Google Sheets/Excel 2026.

Déroulement

Le TP comporte deux parties. La première (L12.1-L12.3) construit le modèle du flux d'information. La seconde (L12.4-L12.6) calcule et interprète les indicateurs. La figure L12.1 montre le résultat attendu de la première partie.

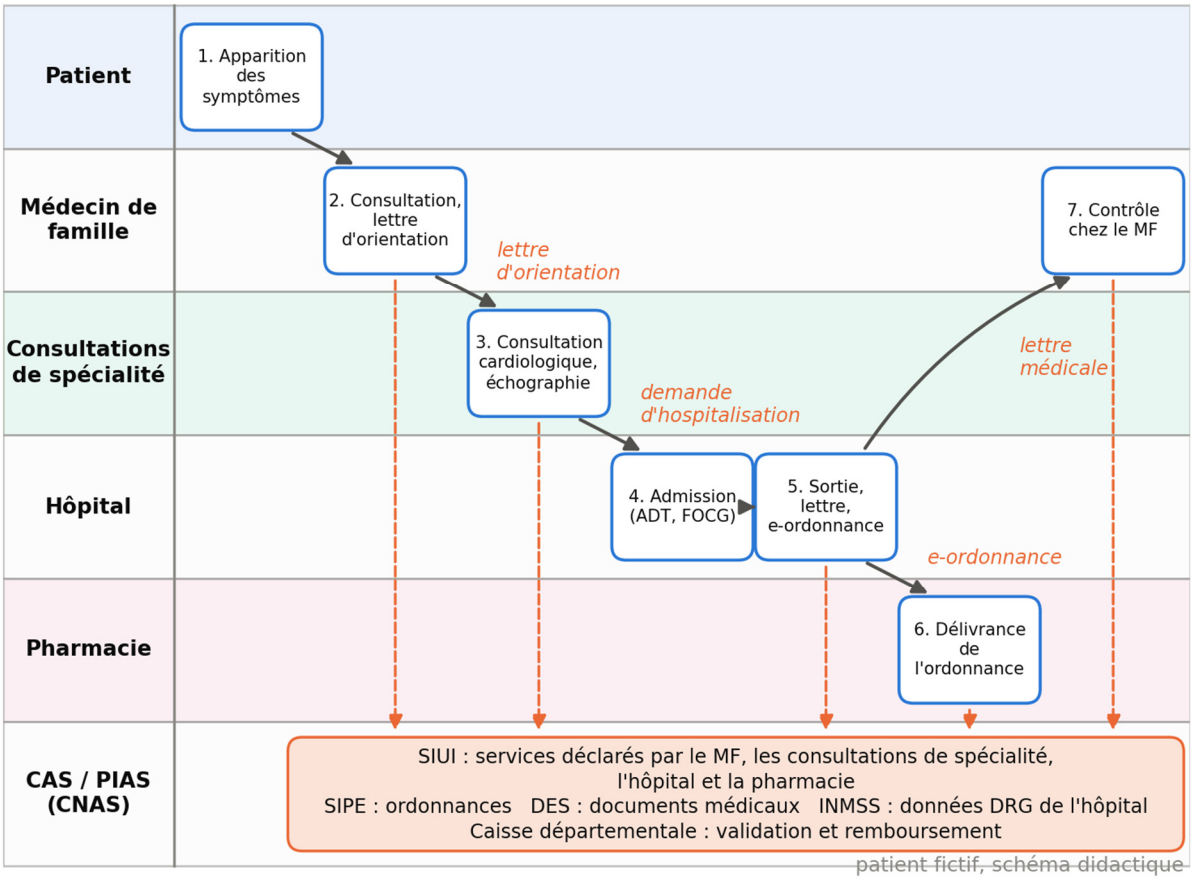


Figure L12.1. Diagramme en couloirs du parcours d'un patient fictif : médecin de famille, consultations externes de spécialité, hôpital, pharmacie et plateforme de la caisse d'assurance maladie. Les flèches pleines indiquent le parcours du patient et les documents qui l'accompagnent, les flèches en pointillés les déclarations vers le SIUI, le SIPE, le DES et l'INMSS. Schéma didactique.

L12.1 Le scénario

Le patient fictif P.X., 64 ans, consulte son médecin de famille pour une dyspnée d'effort apparue au cours du dernier mois. Le médecin l'examine, mesure sa pression artérielle, réalise un électrocardiogramme et l'adresse au cardiologue avec une lettre d'orientation (bilet de trimitere). En consultation externe de spécialité, le cardiologue pratique une échocardiographie, suspecte une insuffisance cardiaque décompensée et établit une demande d'hospitalisation (bilet de internare). Le patient est hospitalisé 5 jours dans le service de cardiologie. À la sortie, il reçoit la lettre médicale et une ordonnance électronique. L'ordonnance est délivrée en pharmacie et, deux semaines plus tard, le patient revient chez le médecin de famille pour un contrôle. Chaque prestataire déclare ses services à la plateforme de la caisse d'assurance maladie, et l'hôpital transmet en outre l'ensemble minimal de données destiné au groupage DRG.

La modélisation des processus (*process mapping*) aide les équipes à comprendre un système complexe avant de le modifier. Une revue systématique de 105 études a cependant constaté qu'aucune ne respectait l'ensemble des dix critères de qualité proposés par les auteurs et que seules 15 % des études avaient associé des patients à l'élaboration de la cartographie [585]. Un langage largement adopté est BPMN (*Business Process Model and Notation*), qui présente des avantages pour la standardisation des flux, mais aussi des limites dans les environnements cliniques très complexes [586]. Le diagramme en couloirs utilisé ici en est une forme simplifiée, suffisante pour un premier modèle.

L12.2 Le diagramme en couloirs dans draw.io

1. Ouvrez https://app.diagrams.net. Lorsque le programme demande où enregistrer, choisissez *Device* (l'ordinateur personnel). Créez un nouveau diagramme et, dans la collection de modèles *Basic*, choisissez *Swimlane Diagram* [582]. Enregistrez immédiatement le fichier sous le nom `L12_Nume_Prenume.drawio`.
2. Le modèle contient un conteneur (*pool*) comportant quelques couloirs (*lanes*). Pour ajouter des couloirs, sélectionnez un couloir et appuyez sur Alt+Entrée, ou utilisez, dans l'onglet *Arrange*, les commandes *Insert Row Above* ou *Insert Row After* [582]. Il vous faut six couloirs.
3. Renommez les couloirs par un double clic sur le titre de chacun : Patient, Médecin de famille, Consultations de spécialité, Hôpital, Pharmacie, CAS/PIAS [582].
4. Faites glisser depuis la bibliothèque *General* un rectangle arrondi pour chaque étape et placez-le dans le couloir de l'acteur qui l'exécute. Numérotez les étapes de 1 à 7, comme dans la figure L12.1. Une étape appartient au couloir de celui qui accomplit l'action, et non de celui qui la demande.
5. Reliez les étapes par des flèches. Passez la souris sur une forme, tirez un connecteur depuis l'une des flèches de direction et déposez-le sur la forme cible [582]. Sur chaque flèche, écrivez par un double clic le document qui circule : lettre d'orientation, demande d'hospitalisation, e-ordonnance, lettre médicale.
6. Ajoutez des flèches en pointillés depuis les étapes 2, 3, 5, 6 et 7 vers le couloir CAS/PIAS. Le style pointillé se choisit dans le panneau *Style*, à l'option de ligne. Écrivez sur ces flèches le système qui reçoit les données : SIUI, SIPE, DES ou INMSS (chapitre 12, tableau 12.2).
7. Vérifiez le diagramme selon trois règles. Chaque étape a un seul acteur. Chaque flèche a un sens. Aucune information n'apparaît "de nulle part" : tout document utilisé à une étape a été produit à une étape antérieure.
8. Exportez le diagramme par *File* > *Export as* > *PNG*, avec un zoom de 200 %, et conservez aussi le fichier `.drawio`, qui reste modifiable.

L12.3 Documents et éléments de données

Pour le scénario de L12.1, remplissez un tableau semblable au tableau L12.1. Les premières lignes sont données comme modèle ; l'étudiant les vérifie et complète le reste. La colonne "Éléments de données minimaux" ne contient que ce qui est nécessaire à l'étape suivante ou à la déclaration. Par exemple, la pharmacie n'a pas besoin de l'échocardiographie, mais elle a besoin du code de diagnostic qui justifie le remboursement du médicament.

Tableau L12.1. Les documents et les éléments de données minimaux à chaque étape du parcours (patient fictif). Source : construction didactique, l'étape 5 d'après [139,465].

Étape	Acteur et système	Document émis	Éléments de données minimaux
2. Consultation MF	médecin de famille, déclaration SIUI	fiche de consultation, lettre d'orientation	identification du patient, date, diagnostic présumé codé, spécialité demandée, médecin émetteur
3. Consultation cardiologique	consultations externes, déclaration SIUI	résultat de l'échocardiographie, demande d'hospitalisation	diagnostic, examens réalisés, motif de l'hospitalisation
4. Admission	hôpital, module ADT, FOCG	FOCG ouverte	date et heure d'admission, service, lit, type d'hospitalisation, diagnostic d'admission
5. Sortie	hôpital, déclaration SIUI, INMSS, SIPE, DES	lettre médicale, e-ordonnance, ensemble minimal de données	diagnostic principal et diagnostics secondaires (ICD-10-AM), actes, date de sortie, état à la sortie
6. Délivrance de l'ordonnance	pharmacie, SIPE	ordonnance délivrée	médicament, quantité, date de délivrance
7. Contrôle MF	médecin de famille	lettre médicale jointe à la fiche	recommandations de sortie, traitement en cours

À l'étape 5, les diagnostics sont codés en ICD-10-AM, 3e édition, et le groupage en DRG se fait avec la classification RO DRG v.1.1 [139]. L'ensemble minimal de données au niveau du patient est collecté par l'INMSS auprès de tous les hôpitaux [465].

L12.4 Le jeu de données et le tableur

Le service fictif de cardiologie compte 4 lits. En avril 2026 (30 jours), il a enregistré 16 sorties, tous les patients ayant été admis et sortis au cours du même mois. Les poids relatifs des quatre groupes DRG sont **fictifs** et choisis à des fins didactiques ; ce ne sont pas les valeurs officielles RO DRG.

Tableau L12.2. Le jeu de données fictif du service (avril 2026). Poids DRG fictifs : A = 0,85 (douleur thoracique sans complications), B = 1,20 (insuffisance cardiaque), C = 1,65 (infarctus du myocarde), D = 2,40 (implantation d'un stimulateur cardiaque). Données fictives, illustratives.

ID	Admission	Sortie	Groupe DRG	Réhospitalisation dans les 30 jours
P01	01.04.2026	07.04.2026	B	Non
P02	01.04.2026	04.04.2026	A	Non
P03	02.04.2026	12.04.2026	C	Oui, non programmée
P04	03.04.2026	08.04.2026	B	Non
P05	04.04.2026	06.04.2026	A	Non
P06	06.04.2026	15.04.2026	D	Non
P07	08.04.2026	14.04.2026	B	Non
P08	09.04.2026	12.04.2026	A	Non
P09	12.04.2026	19.04.2026	C	Non
P10	13.04.2026	17.04.2026	B	Oui, programmée
P11	15.04.2026	23.04.2026	D	Non
P12	16.04.2026	18.04.2026	A	Non
P13	18.04.2026	26.04.2026	C	Non
P14	20.04.2026	25.04.2026	B	Non
P15	23.04.2026	29.04.2026	B	Non
P16	24.04.2026	30.04.2026	C	Non

Organisez le tableur comme suit :

1. Dans la ligne 1, écrivez les en-têtes : A = ID, B = Admission, C = Sortie, D = Jours, E = DRG, F = Poids, G = Réhospitalisation (Oui/Non), H = Type de réhospitalisation (programmée, non programmée ou vide). Saisissez les 16 enregistrements dans les lignes 2 à 17. Les dates s'écrivent dans le format régional du tableur (pour la Roumanie, jj.mm.aaaa).
2. Dans une zone séparée, construisez la table des poids : M1 = DRG, N1 = Poids, O1 = Cas, P1 = DMS par groupe. Dans M2:M5, écrivez A, B, C, D, et dans N2:N5 les poids 0,85, 1,20, 1,65 et 2,40. Dans K1, écrivez "Lits" et dans K2 la valeur 4. Dans K3, écrivez 30, le nombre de jours du mois.
3. Les journées d'hospitalisation : dans D2, écrivez `=C2-B2` et recopiez la formule jusqu'à D17. Le jour de sortie n'est pas compté : Eurostat calcule les journées d'hospitalisation comme la différence entre la date de sortie et la date d'admission [474]. Si le résultat s'affiche sous forme de date, changez le format de la cellule en nombre.
4. Le poids de chaque cas : dans F2, écrivez `=VLOOKUP(E2;$M$2:$N$5;2;0)` et recopiez jusqu'à F17. Le dernier argument, 0, demande une correspondance exacte et fonctionne quelle que soit la langue de l'interface (il équivaut à FALSE). Le symbole `$` fige la table des poids lors de la recopie.
5. Le nombre de cas par groupe : dans O2, écrivez `=COUNTIF($E$2:$E$17;M2)` et recopiez jusqu'à O5. La somme O2:O5 doit être égale à 16.
6. La DMS par groupe : dans P2, écrivez `=AVERAGEIF($E$2:$E$17;M2;$D$2:$D$17)` et recopiez jusqu'à P5.

Dans les tableurs configurés pour la Roumanie (comme pour la France), les arguments sont séparés par un point-virgule et les décimales par une virgule. Avec des paramètres en anglais, le séparateur est la virgule. Dans Excel avec une interface en français, les noms des fonctions peuvent apparaître traduits (par exemple RECHERCHEV, NB.SI, MOYENNE.SI, SOMMEPROD).

L12.5 Le calcul des indicateurs

Écrivez les formules ci-dessous dans une zone de résultats (par exemple, la colonne R) et comparez le résultat avec les valeurs de contrôle du tableau L12.3.

1. **La durée moyenne de séjour** est le nombre total de journées d'hospitalisation divisé par le nombre de sorties [474] : `=SUM(D2:D17)/COUNT(D2:D17)`.
2. **Le taux d'occupation des lits** est le rapport entre les journées d'hospitalisation réalisées et les journées-lits disponibles, multiplié par 100 : `=SUM(D2:D17)/(K2*K3)*100`.
3. **L'indice de case-mix** est la moyenne des poids relatifs des cas sortis. Calculez-le de deux façons, qui doivent donner le même résultat : `=AVERAGE(F2:F17)` et `=SUMPRODUCT(N2:N5;O2:O5)/SUM(O2:O5)` [584].
4. **Le taux de réhospitalisation à 30 jours**, global : `=COUNTIF(G2:G17;"Oui")/COUNTA(A2:A17)*100`.
5. **Le taux de réhospitalisation non programmée** : `=COUNTIFS(G2:G17;"Oui";H2:H17;"non programmée")/COUNTA(A2:A17)*100` [583].
6. Construisez un graphique en colonnes de la DMS par groupe (P2:P5), avec les groupes DRG sur l'axe horizontal, et ajoutez-le au fichier rendu.

Tableau L12.3. Les indicateurs calculés et les valeurs de contrôle (données fictives).

Indicateur	Formule (Google Sheets)	Valeur de contrôle
Journées d'hospitalisation, total	=SUM(D2:D17)	90
Durée moyenne de séjour	=SUM(D2:D17)/COUNT(D2:D17)	5,625 jours (5,6)
Taux d'occupation	=SUM(D2:D17)/(K2*K3*)100	75 %
Indice de case-mix	=AVERAGE(F2:F17)	1,375
Réhospitalisations à 30 jours, total	=COUNTIF(G2:G17;"Oui")/COUNTA(A2:A17)*100	12,5 %
Réhospitalisations non programmées	=COUNTIFS(G2:G17;"Oui";H2:H17;"non programmée")/COUNTA(A2:A17)*100	6,25 %
DMS par groupe A, B, C, D	=AVERAGEIF(...)	2,5 / 5,33 / 7,75 / 8,5 jours

Si un résultat diffère, vérifiez d'abord le format des dates et les références absolues. Une date saisie comme texte (par exemple, avec une espace finale) produit une erreur lors de la soustraction. Autres erreurs fréquentes : compter le jour de sortie, ce qui donnerait 106 jours au lieu de 90 et une DMS de 6,6 jours, diviser par le nombre de lits sans multiplier par le nombre de jours du mois, et écrire "Oui " avec une espace, que la fonction COUNTIF ne reconnaît plus. Pour chaque formule, calculez un cas à la main et comparez-le avec la cellule.

L12.6 Interprétation

Répondez par écrit, en 8 à 10 lignes, aux questions suivantes :

1. Un taux d'occupation de 75 % est-il élevé ou faible pour un service de cardiologie ? Quelles informations faudrait-il encore pour décider si le nombre de lits est adapté ?
2. L'ICM est de 1,375. Que signifie le fait qu'il soit supérieur à 1, et qu'arriverait-il à l'indice si le service admettait, le mois suivant, davantage de cas du groupe A ?
3. La DMS augmente du groupe A au groupe D. Est-ce une observation attendue ? Que signalerait une DMS élevée dans un groupe de faible poids ?
4. Le taux global de réhospitalisation est de 12,5 %, mais le taux de réhospitalisation non programmée est de 6,25 %. Lequel des deux dit quelque chose de la qualité des soins ? Une étude portant sur 8 878 patients de neurochirurgie a conclu que le taux de réhospitalisation à 30 jours, pris isolément, n'est pas une mesure adéquate de la qualité, car il n'a pas de définition homogène et nécessite un ajustement sur le risque [477].
5. Pourquoi ne peut-on pas comparer la DMS de ce service, calculée sur un mois et 16 cas, avec l'intervalle de 5,2 à 9,4 jours rapporté pour la plupart des États de l'UE en 2023 [474] ?
6. Sur le diagramme de L12.2, indiquez l'étape à laquelle une erreur de codage du diagnostic modifierait la valeur de l'ICM calculée par l'hôpital.

Travail à rendre

Chaque étudiant dépose sur la plateforme du cours un seul fichier d'archive nommé `L12_Nume_Prenume.zip`, qui contient :

- le diagramme en couloirs, aux formats `.drawio` et `.png` (30 points : 6 couloirs correctement nommés, 7 étapes placées dans le bon couloir, les documents écrits sur les flèches, les déclarations vers CAS/PIAS).

- le tableau des documents et des éléments de données, rempli pour toutes les étapes (20 points).
- le tableur avec les formules visibles (et pas seulement les valeurs) et le graphique de la DMS par groupe (30 points : 5 points pour chaque indicateur correctement calculé et 5 points pour le graphique).
- les réponses aux questions d'interprétation (20 points).

À l'examen pratique, l'étudiant peut recevoir un autre jeu de données fictif (un autre nombre de lits ou d'autres poids) et doit recalculer les indicateurs en 15 minutes au plus.

Variante hors ligne

Si la salle de TP n'a pas accès à Internet, le diagramme se dessine dans l'application de bureau draw.io, installée au préalable depuis la page officielle des versions [581]. L'application de bureau enregistre les fichiers localement et fonctionne sans connexion. Si elle n'est pas disponible non plus, le diagramme se dessine sur une feuille A4 posée à l'horizontale, avec six bandes tracées à la règle, et l'étudiant la photographie pour la rendre.

Les indicateurs se calculent dans n'importe quel tableur installé localement, avec les mêmes formules. Faute d'ordinateur, le calcul se fait à la main sur le tableau L12.2. On additionne les 16 durées de séjour (90 jours), on divise par 16 pour la DMS, par 4 x 30 = 120 journées-lits pour le taux d'occupation, et, pour l'ICM, on additionne les poids (4 x 0,85 + 6 x 1,20 + 4 x 1,65 + 2 x 2,40 = 22,00) avant de diviser par 16.

Travaux pratiques 13. L'évaluation critique d'un outil d'IA : de la matrice de confusion au débat

Objectif

L'étudiant calcule les indicateurs de performance d'un outil d'IA fictif, observe comment les valeurs prédictives varient avec la prévalence, construit à la main une courbe ROC, applique une courte liste de vérification dérivée de TRIPOD+AI à un article réel et soutient, dans un débat avec des rôles attribués, une position argumentée sur la responsabilité liée à un compte rendu d'imagerie. Le TP applique les notions du chapitre 13 (sections 13.3-13.7). Le protocole de calcul des indicateurs diagnostiques de [6] est repris et mis à jour pour Google Sheets/Excel 2026.

Outils

- **Google Sheets**, application web gratuite accessible avec un compte Google, consultée le 27 septembre 2026 (https://sheets.google.com). On utilise les fonctions `SUM`, `COUNTIF`, `ROUND`, `MAX` et `SUMPRODUCT` [583] ainsi que les graphiques à nuage de points (*scatter*) et en courbes (*line chart*) [587].
- **LibreOffice Calc 26.8.0**, alternative gratuite, installée localement, pour la variante hors ligne [588]. Microsoft Excel peut être utilisé s'il est disponible, avec les mêmes formules.
- **L'article à analyser** : Wong et ses collaborateurs, validation prospective multicentrique de la deuxième version du modèle de prédiction du sepsis d'Epic, JAMA Network Open 2026, en libre accès [503].
- **La liste TRIPOD+AI** (BMJ 2024), sous la forme abrégée du tableau L13.3 [505].
- Calculatrice de poche, papier millimétré (pour la variante hors ligne).

Toutes les données des parties A et B sont fictives, construites pour l'exercice. Aucune donnée de patient n'est utilisée.

Déroulement

L13.1 Partie A. La matrice de confusion et l'effet de la prévalence (35 min)

L'outil fictif "PneumoScan" lit des radiographies thoraciques et signale les pneumonies. Il a été testé sur 1 000 radiographies, dont 100 présentaient une pneumonie confirmée par le standard de référence (prévalence de 10 %). Les résultats figurent dans le tableau L13.1.

Tableau L13.1. La matrice de confusion de l'outil fictif PneumoScan (n = 1 000, prévalence de 10 %) et les formules pour Google Sheets, avec les cellules B2 = VP, C2 = FP, B3 = FN, C3 = VN. Données fictives. Formules d'après [499,500].

	Pneumonie présente	**Pneumonie absente**	**Indicateur**	**Formule dans Sheets**
PneumoScan positif	VP = 85	FP = 90	Sensibilité	=B2/(B2+B3)
PneumoScan négatif	FN = 15	VN = 810	Spécificité	=C3/(C3+C2)
			VPP	=B2/(B2+C2)
			VPN	=C3/(C3+B3)
			Exactitude	=(B2+C3)/SUM(B2:C3)
			Score F1	=2*B2/(2*B2+C2+B3)

Étapes :

1. Ouvrez une nouvelle feuille et nommez-la `L13_NumePrenume`. Dans les cellules A1:C3, construisez la matrice : la ligne 1 avec les en-têtes "Pneumonie présente" et "Pneumonie absente", la colonne A avec "Positif" et "Négatif", puis les valeurs 85, 90, 15, 810 dans B2, C2, B3, C3.
2. Dans la colonne E, écrivez le nom des indicateurs, et dans la colonne F les formules du tableau L13.1. Mettez la colonne F au format pourcentage avec une décimale.
3. Vérifiez les résultats : Se = 85,0 %, Sp = 90,0 %, VPP = 48,6 %, VPN = 98,2 %, exactitude = 89,5 %, F1 = 0,618. Si une valeur diffère, vérifiez les adresses des cellules dans la formule.
4. Construisez, dans H1:K13, un tableau (les étiquettes dans la colonne H, les contextes dans I, J, K) pour trois contextes cliniques ayant la même sensibilité (0,85) et la même spécificité (0,90) : **dépistage dans la population générale (2 %)**, **patients ambulatoires présentant une toux (10 %)**, **patients hospitalisés avec fièvre et râles (40 %)**. Pour chaque colonne : le nombre de malades `=ROUND(1000*prévalence)`, VP `=ROUND(malades*0,85)`, FN `=malades-VP`, VN `=ROUND((1000-malades)*0,90)`, FP `=(1000-malades)-VN`, puis les mêmes six indicateurs. Le séparateur décimal des formules dépend des paramètres régionaux de la feuille (virgule pour la Roumanie, point pour les États-Unis).
5. Comparez. À une prévalence de 2 %, vous devriez obtenir une VPP ≈ 14,8 %, à 10 % une VPP ≈ 48,6 %, à 40 % une VPP = 85,0 %, alors que l'exactitude reste comprise entre 88 % et 90 %.
6. Calculez l'exactitude d'un "test" qui déclare tout le monde sain, à une prévalence de 2 % (réponse : 98 %). Écrivez dans la feuille, dans une cellule de commentaire, deux phrases : pourquoi l'exactitude est un indicateur trompeur lorsque la maladie est rare, et pourquoi la VPP de l'outil ne peut pas être "transportée" d'un hôpital de maladies infectieuses vers un programme de dépistage.

Remarque méthodologique : le score F1 ignore les vrais négatifs et, dans une analyse de 32 indicateurs, il n'a satisfait à aucun des deux critères de qualité analysés par les auteurs [500]. On ne le calcule ici que parce qu'il apparaît fréquemment dans les articles d'IA.

L13.2 Partie B. La courbe ROC construite pas à pas (35 min)

Le même type d'outil ne produit pas directement "positif/négatif", mais un score compris entre 0 et 1. Le tableau L13.2 contient les scores de 16 patients fictifs, 6 atteints de la maladie (état = 1) et 10 indemnes (état = 0).

Tableau L13.2. Les scores de l'outil fictif pour 16 patients, classés par ordre décroissant. Données fictives, construites pour l'exercice.

Patient	Score	État	Patient	Score	État
P01	0,95	1	P09	0,41	0
P02	0,88	1	P10	0,34	1
P03	0,81	1	P11	0,30	0
P04	0,76	0	P12	0,26	0
P05	0,67	1	P13	0,19	0
P06	0,62	0	P14	0,12	0
P07	0,58	1	P15	0,08	0
P08	0,49	0	P16	0,05	0

Étapes :

1. Dans une nouvelle feuille, copiez les données dans les colonnes A (patient), B (score), C (état), à partir de la ligne 3. Laissez la ligne 2 pour le point de départ de la courbe (0, 0). Vérifiez que le tableau est trié par score décroissant (menu Données, Trier la plage).
2. Chaque ligne représente un seuil : les patients dont le score est supérieur ou égal au score de la ligne considérée sont déclarés positifs. Dans D3, écrivez `=SUM($C$3:C3)` (vrais positifs cumulés), dans E3 `=COUNTIF($C$3:C3,0)` (faux positifs cumulés) et étendez les formules jusqu'à la ligne 18. Dans D2 et E2, écrivez 0. Si la feuille utilise les paramètres régionaux roumains (ou français), les arguments sont séparés par un point-virgule : `=COUNTIF($C$3:C3;0)`.
3. Dans F2, calculez la sensibilité `=D2/6`, dans G2 le taux de faux positifs (1 - spécificité) `=E2/10`, et recopiez jusqu'à la ligne 18. La dernière ligne doit donner F = 1 et G = 1.
4. Sélectionnez F2:G18 et insérez un graphique à nuage de points. Dans l'éditeur du graphique, choisissez la colonne G pour l'axe X et la colonne F comme série (axe Y). Google Sheets ne relie pas les points d'un graphique à nuage de points ; pour obtenir une courbe continue, choisissez le type graphique en courbes, avec la même colonne G sur l'axe X (Excel et LibreOffice Calc proposent directement la variante nuage de points avec lignes). Ajoutez la diagonale (0,0)-(1,1) comme seconde série. La courbe obtenue doit ressembler à celle de la figure L13.1.
5. Calculez l'aire sous la courbe par la méthode des trapèzes : `=SUMPRODUCT((G3:G18-G2:G17)*(F3:F18+F2:F17)/2)`. Le résultat correct est 0,883.
6. Vérifiez d'une autre façon : comptez, pour chacun des 6 malades, combien des 10 sujets sains ont un score inférieur. La somme des paires correctement ordonnées est de 53 sur 60 possibles, et 53/60 = 0,883. L'AUC est donc la probabilité qu'un malade choisi au hasard ait un score plus élevé qu'un sujet sain choisi au hasard [501].
7. Dans H2, calculez l'indice de Youden `=F2-G2`, recopiez vers le bas et trouvez le maximum avec `=MAX(H2:H18)`. Le seuil optimal selon ce critère est 0,58 (Se 0,83, Sp 0,80, J = 0,63). Discutez en groupe des raisons pour lesquelles, pour un outil de triage, on pourrait préférer le seuil 0,34 (Se 1,00, Sp 0,60), même si l'indice de Youden est un peu plus faible.

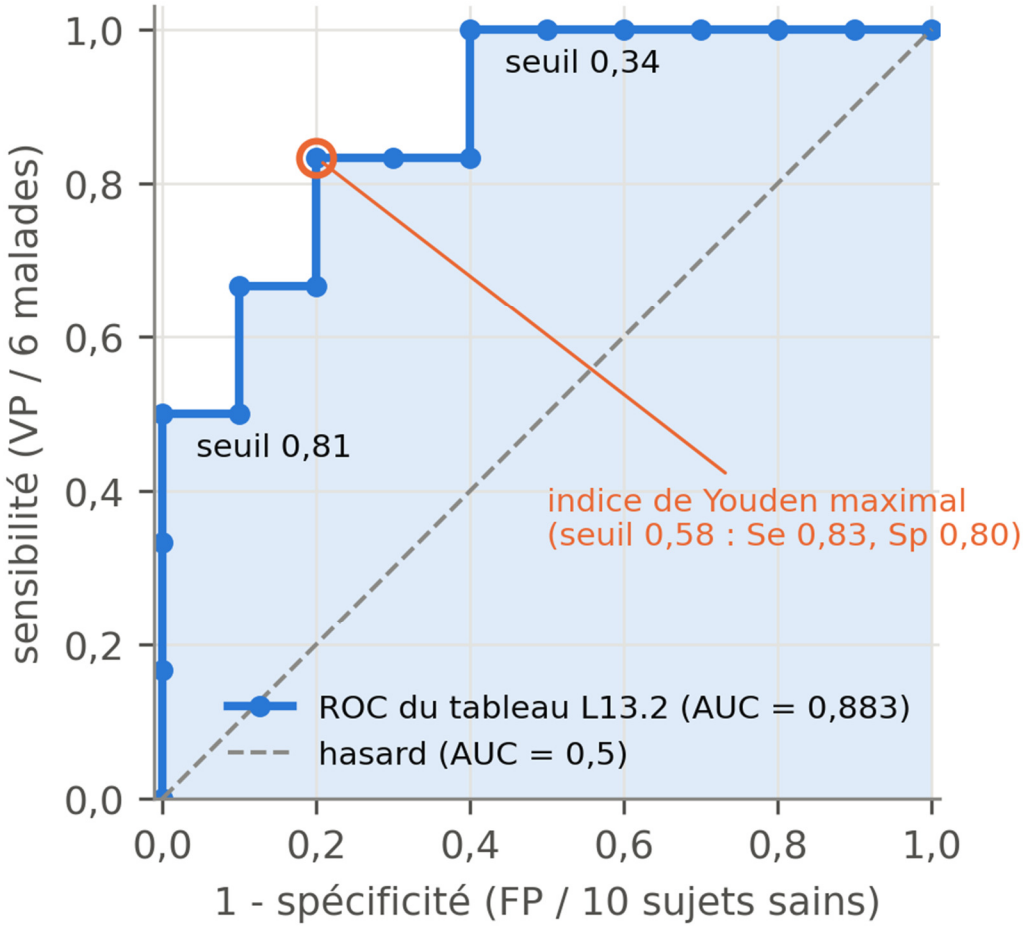


Figure L13.1. La courbe ROC construite à partir du tableau L13.2 : chaque point correspond à un seuil, les marches verticales à un malade, les marches horizontales à un sujet sain. AUC calculée par la méthode des trapèzes = 0,883 ; le point entouré marque l'indice de Youden maximal. Données fictives, calculées dans le script de la figure.

Erreurs fréquentes dans cette partie : le tri croissant au lieu du tri décroissant (la courbe apparaît sous la diagonale), la division par le nombre total de patients au lieu de 6 et de 10 respectivement, et l'omission du point de départ (0, 0), sans lequel la courbe ne part plus de l'origine du graphique. Si deux patients avaient le même score, l'un malade et l'autre sain, le seuil les déclarerait positifs simultanément, et la courbe formerait un segment oblique au lieu d'une marche. Le tableau L13.2 ne contient aucun score identique, précisément pour que les étapes soient faciles à suivre.

L13.3 Partie C. L'évaluation critique d'une étude publiée (40 min)

L'article choisi valide prospectivement, dans quatre systèmes de santé des États-Unis, la deuxième version d'un modèle commercial de prédiction du sepsis. Le résumé rapporte 227 091 hospitalisations, dont 7 401 (3,3 %) répondaient aux critères Sepsis-3, une AUC au niveau de l'hospitalisation comprise entre 0,82 et 0,92 selon le centre et une VPP comprise entre 0,13 et 0,26 [503]. La version précédente du même modèle avait obtenu, lors de la validation externe de 2021, une AUC de 0,63 [502].

Travaillez en binôme. Lisez le texte intégral et remplissez la liste du tableau L13.3, en notant pour chaque élément "oui", "partiellement" ou "non" ainsi que la page ou la section où vous avez trouvé l'information. La liste est une sélection didactique des thèmes de TRIPOD+AI ; elle ne remplace pas la liste officielle de 27 éléments [505].

Tableau L13.3. Courte liste de vérification pour une étude de validation d'un modèle prédictif, dérivée des thèmes de TRIPOD+AI. Source : adapté d'après [505].

N°	Question	Où chercher
1	Le titre et le résumé indiquent-ils qu'il s'agit d'une étude de validation d'un modèle prédictif ?	Titre, résumé
2	La source des données, les centres et la période sont-ils décrits ?	Méthodes
3	Le résultat prédit (le sepsis) est-il défini explicitement et de la même manière dans tous les centres ?	Méthodes
4	Le modèle est-il décrit (version, moment de la prédiction, horizon temporel) ?	Méthodes
5	Le nombre de patients et le nombre d'événements sont-ils rapportés ?	Résultats
6	La discrimination est-elle rapportée, avec son intervalle de confiance ?	Résultats
7	La calibration est-elle rapportée (graphique ou indicateurs) ?	Résultats, annexes
8	Des indicateurs d'utilité clinique (VPP, charge d'alertes) sont-ils rapportés ?	Résultats
9	La performance est-elle analysée par sous-groupes (sexe, âge, origine ethnique, centre) ?	Résultats, annexes
10	Les limites, le financement, les conflits d'intérêts et la disponibilité du code sont-ils déclarés ?	Discussion, déclarations

Faites ensuite deux calculs :

1. Avec une VPP = 0,13, combien d'alertes sur 100 sont de fausses alarmes ? (87.) Avec une VPP = 0,26 ? (74.)
2. Le nombre de patients à évaluer pour trouver un cas réel (*number needed to evaluate*) est égal à 1/VPP. Calculez-le pour les deux extrêmes (environ 8 et 4).

Terminez par un paragraphe de 80 à 120 mots qui répond à la question : "Recommanderiez-vous l'activation de ce modèle dans un hôpital de Roumanie et, si oui, à quelles conditions de surveillance ?". Utilisez les notions de validation externe, de prévalence et de dérive des données du chapitre 13. Pour une évaluation formelle du risque de biais, il existe l'outil PROBAST+AI [506], facultatif pour les étudiants intéressés.

L13.4 Partie D. Débat structuré (40 min)

La motion : "Un système d'IA devrait pouvoir signer seul un compte rendu d'imagerie."

Rôles. Le groupe se divise en cinq équipes de 2 ou 3 étudiants, chacune avec un rôle, auxquelles s'ajoutent un modérateur (l'enseignant ou un étudiant) et le public :

- **le clinicien (radiologue)** : la charge de travail, les erreurs humaines, l'expérience des systèmes d'assistance
- **le patient** : le droit de savoir qui a lu l'examen, la confiance, la voie de réclamation
- **le développeur** : la performance rapportée, le coût, la mise à jour du modèle
- **le régulateur** : la classification du logiciel comme dispositif médical et les obligations applicables aux systèmes à haut risque
- **l'éthicien** : la responsabilité, l'équité, le biais d'automatisation (*automation bias*).

Règles. Chaque équipe dispose de 3 minutes pour son exposé d'ouverture ; suivent un tour de répliques de 1 minute chacune et 5 minutes de questions du public. Tout argument factuel doit être accompagné d'une source tirée du chapitre 13 ou de la liste ci-dessous. Les affirmations sans source ne sont pas notées. À la fin, le public vote deux fois, avant et après le débat, et le modérateur note le changement.

Sources minimales recommandées : les obligations de l'AI Act applicables aux systèmes à haut risque, dont le contrôle humain [523], la règle 11 du MDR pour la classification des logiciels [284], l'étude MASAI sur la lecture des mammographies assistée par l'IA [271], l'expérience sur le biais d'automatisation chez les radiologues [281] et la position de l'ICMJE, selon laquelle un outil d'IA ne peut pas être auteur, car il ne peut pas répondre du travail [202].

Grille d'évaluation par équipe (10 points) : clarté et cohérence de l'argumentation (0-3), utilisation correcte des preuves, avec chiffres et sources (0-3), réponse aux contre-arguments (0-2), respect du rôle et du temps (0-1), conclusion formulée en une phrase (0-1).

À la fin du débat, le modérateur résume en cinq minutes les points d'accord et de désaccord. Le résumé doit séparer trois questions que l'on confond facilement : le système lit-il suffisamment bien (preuves de performance), qui répond de ses erreurs (responsabilité juridique et déontologique) et que permet la loi en vigueur au 27 septembre 2026 (réglementation). Chaque équipe note sur sa fiche l'argument de la partie adverse qu'elle juge le plus difficile à réfuter.

Travail à rendre

On dépose sur la plateforme de la discipline un seul fichier (Google Sheets partagé avec l'enseignant ou fichier .xlsx/.ods), comportant quatre feuilles :

1. **Partie A** : la matrice, les six indicateurs et le tableau pour les trois prévalences, avec les deux phrases d'interprétation (3 points).
2. **Partie B** : le tableau cumulatif, le graphique ROC, l'AUC par la méthode des trapèzes, la vérification par paires et l'indice de Youden (3 points).
3. **Partie C** : la liste du tableau L13.3 remplie, les deux calculs et le paragraphe de recommandation (2 points).
4. **Partie D** : la fiche de l'équipe, avec les arguments et les sources utilisés, ainsi que la note obtenue dans la grille, convertie (2 points).

Barème indicatif à l'examen pratique : l'étudiant reçoit une autre matrice de confusion et une autre prévalence et doit calculer en 10 minutes la Se, la Sp, la VPP, la VPN et l'exactitude, puis expliquer pourquoi la VPP a changé.

Variante hors ligne

En l'absence d'accès à Internet, les parties A et B se font dans LibreOffice Calc 26.8.0 installé localement, avec les mêmes formules (dans la version roumaine ou française du programme, le séparateur d'arguments peut être le point-virgule). Si aucun ordinateur n'est disponible, on travaille sur papier : la matrice et les indicateurs avec la calculatrice de poche, la courbe ROC sur papier millimétré, avec 10 carreaux à l'horizontale (un sujet sain = un pas vers la droite) et 6 à la verticale (un malade = un pas vers le haut), et l'AUC en comptant les carreaux situés sous la courbe (53 sur 60). Pour la partie C, l'enseignant distribue le texte imprimé de l'article. Le débat ne dépend pas du réseau.

Travaux pratiques 14. L'examen pratique : modèle de sujet, barème et révision

Objectif

Cette séance prépare l'examen pratique. Conformément au plan d'enseignement de la discipline, l'examen pratique représente 30 % de la note finale. À la fin de la séance, l'étudiant connaît le format de l'épreuve, a traité un sujet type complet, sait comment chaque point est attribué et dispose d'un plan de révision pour les TP 1-13. Le sujet ci-dessous est un modèle didactique proposé par cet ouvrage : le nombre d'exercices, la durée et le barème peuvent être fixés différemment par le titulaire de la discipline, et la version officielle est annoncée en cours. Tous les cas, valeurs et fichiers de cette séance sont fictifs.

Outils

Les outils sont ceux des TP précédents, tous gratuits, avec des adresses vérifiées le 27 septembre 2026 :

- Google Forms et Google Sheets (ou LibreOffice Calc), pour les exercices S1 et S3 [123,124].
- Le navigateur CIM-10 de l'OMS, version 2019, le navigateur CIM-11 MMS, version 2026-01, et le guide de référence de la CIM-11, pour l'exercice S2 [133,134,141].
- Fiji, dans la version installée au laboratoire, avec un fichier DICOM anonymisé et l'image d'exemple Blobs, pour l'exercice S4 [245,557].
- PubMed, pour la vérification des références de l'exercice S6 [41].
- Une calculatrice de poche, du papier et un stylo, pour la variante hors ligne.

L'étudiant n'utilise pas de chatbot pendant l'épreuve. L'exercice S6 porte sur une réponse d'un modèle de langage imprimée à l'avance par l'enseignant.

Déroulement

Le modèle proposé comporte six exercices, dure 90 minutes et est noté sur 10 points (figure L14.1). Les exercices peuvent être traités dans n'importe quel ordre. Chaque réponse est enregistrée dans un document de remise unique, nommé `L14_Nume_Prenume`, et les fichiers de travail (la feuille de calcul, les captures d'écran sans données personnelles) sont joints à la fin.

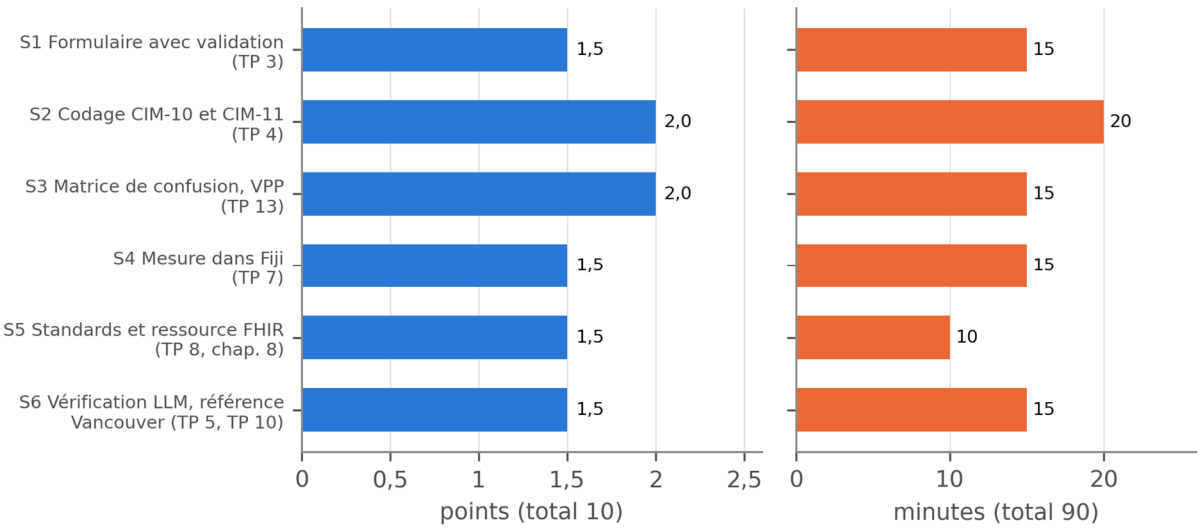


Figure L14.1. Le modèle d'examen pratique proposé dans l'ouvrage : le barème (à gauche) et la durée indicative (à droite) pour chacun des six exercices, avec le TP dont provient chacun d'eux. Modèle didactique, et non règlement.

L14.1 Le sujet type

S1. Formulaire avec validation (1,5 point, 15 minutes, TP 3). L'étude fictive "Le sommeil et la consommation de café chez les étudiants de première année" recueille : le code pseudonyme, l'âge, les heures de sommeil de la nuit précédente, la consommation de café (oui/non) et, uniquement pour ceux qui boivent du café, le nombre de tasses par jour. Consignes :

1. Rédigez le dictionnaire de données des cinq variables : nom court, libellé, type, unité, valeurs autorisées.
2. Construisez le formulaire dans Google Forms, avec la validation du code par l'expression régulière `^S[0-9]{3}$`, l'âge compris entre 18 et 40 ans et les heures de sommeil entre 0 et 14 [123].
3. Configurez le saut vers la section du nombre de tasses uniquement pour la réponse "oui" [124].
4. Expliquez en une phrase pourquoi le formulaire ne demande pas l'adresse e-mail.

S2. Codage en CIM-10 et en CIM-11 (2 points, 20 minutes, TP 4). Codez l'affection principale de chaque cas fictif en CIM-10 (version OMS 2019) et en CIM-11 (MMS 2026-01). Notez la version à côté de chaque code.

1. Femme, 64 ans, consultation de contrôle pour un diabète de type 2, sans complication documentée.
2. Femme, 30 ans, cystite aiguë.
3. Femme, 82 ans, a trébuché sur un tapis à son domicile et est tombée. Diagnostic : fracture sous-capitale fermée du col du fémur gauche.

S3. Performance d'un outil d'IA (2 points, 15 minutes, TP 13). L'outil fictif "OtoScan" a été testé sur 1 000 patients, dont 100 étaient malades selon le standard de référence. Résultats : VP = 80, FP = 45, FN = 20, VN = 855.

1. Calculez la sensibilité, la spécificité, la valeur prédictive positive (VPP), la valeur prédictive négative (VPN) et l'exactitude [499].
2. Recalculez la VPP et la VPN pour un programme de dépistage où la prévalence est de 2 %, en conservant la sensibilité et la spécificité.

3. Expliquez en deux phrases pourquoi la VPP a changé.

S4. Mesure dans Fiji (1,5 point, 15 minutes, TP 7). Ouvrez le fichier DICOM anonymisé que vous avez reçu (coupe TDM fictive, didactique).

1. Lisez dans l'en-tête Pixel Spacing (0028,0030), Window Center (0028,1050), Rescale Slope (0028,1053) et Rescale Intercept (0028,1052) [245].
2. Un pixel a la valeur stockée 1 064, avec Slope = 1 et Intercept = -1 024. Calculez la valeur en unités Hounsfield.
3. Pixel Spacing vaut 0,78 mm. Une ligne tracée sur le diamètre d'une structure mesure 64 pixels. Calculez la longueur en millimètres et vérifiez-la avec Analyze > Measure.
4. Sur l'image Blobs, appliquez le seuil Default, puis Analyze Particles avec la taille 50-Infinity et Exclude on edges. Notez le nombre d'objets et expliquez ce que fait l'option Exclude on edges.

S5. Standards d'interopérabilité (1,5 point, 10 minutes, TP 8, chapitre 8). La patiente fictive A.T. est hospitalisée. Pour chaque étape, nommez le standard ou le composant principal :

1. l'analyseur envoie la glycémie au système du laboratoire et au système de l'hôpital, avec un code pour l'analyse
2. le scanner envoie la série à l'archive d'images
3. une application mobile autorisée lit la pression artérielle dans le dossier électronique au moyen d'une API
4. un médecin d'un autre État membre demande le résumé de la patiente
5. le médecin signe la lettre médicale de sorte que la signature ait l'effet juridique d'une signature manuscrite.

Ensuite, dans le fragment JSON ci-dessous, indiquez le patient, le paramètre mesuré, la valeur, l'unité et le moment de la mesure [567].

```
{"resourceType": "Observation", "status": "final",
 "code": {"coding": [{"system": "http://loinc.org", "code": "8867-4"}]},
 "subject": {"reference": "Patient/demo-17"},
 "effectiveDateTime": "2026-10-12T08:30:00+03:00",
 "valueQuantity": {"value": 88, "unit": "/min"}}
```

S6. Vérification d'une réponse générée et citation (1,5 point, 15 minutes, TP 5, TP 10). Vous recevez la réponse imprimée d'un chatbot pour le cas fictif "homme de 45 ans présentant une céphalée d'installation brutale pendant un effort, la plus violente de sa vie, accompagnée de nausées, qui demande quel antalgique prendre". La réponse contient quatre affirmations numérotées et deux références avec PMID.

1. Classez chaque affirmation comme confirmée, infirmée ou invérifiable et indiquez la catégorie d'erreur selon la grille du TP 10 [406].
2. Recherchez les deux PMID dans PubMed et notez si la référence existe et si elle étaye l'affirmation.
3. Rédigez selon le style Vancouver la référence de l'article ayant les métadonnées suivantes : auteurs Walters WH et Wilder EI, titre "Fabrication and errors in the bibliographic citations generated by ChatGPT", revue Scientific Reports, année 2023, volume 13, numéro 1, article 14045, DOI 10.1038/s41598-023-41032-5 [201].
4. Rédigez, en deux phrases, la déclaration d'utilisation de l'IA pour un rapport qui citerait cette réponse, conformément aux recommandations de l'ICMJE [202].

L14.2 Corrigés et barème

Les corrigés ci-dessous sont ceux que l'enseignant utilise pour la correction. Le tableau L14.1 résume le barème.

S1. Le dictionnaire de données contient, par exemple : `cod` (texte, S suivi de trois chiffres), `varsta` (numérique entier, années, 18-40), `somn_ore` (numérique, heures, 0-14), `cafea` (nominal, oui/non), `cesti_zi` (numérique entier, renseigné uniquement si `cafea` = oui). L'adresse e-mail n'est pas demandée, car c'est un identifiant direct, inutile pour l'objectif de l'étude, et les données concernant la santé ne sont collectées que dans la mesure nécessaire.

S2. La résolution suit le tableau L4.2 [133,134,141] :

- cas 1 : E11.9 en CIM-10 et 5A11 en CIM-11
- cas 2 : N30.0 en CIM-10 et le cluster GC00.1&XT5R en CIM-11
- cas 3 : S72.0, avec W01 comme cause externe, en CIM-10 et le cluster NC72.20&XK8G&XJ44E/PA60 en CIM-11. Les extensions de latéralité et de fracture fermée sont reliées par "&" au code de la fracture, et le code de cause externe PA60 est relié par "/".

S3. La sensibilité est de 80/100 = 80,0 %, la spécificité de 855/900 = 95,0 %, la VPP de 80/125 = 64,0 %, la VPN de 855/875 = 97,7 % et l'exactitude de 935/1 000 = 93,5 %. À une prévalence de 2 %, il y a 20 malades : VP = 16, FN = 4. Parmi les 980 sujets sains, VN = 931 et FP = 49. On obtient VPP = 16/65 = 24,6 % et VPN = 931/935 = 99,6 %. La VPP diminue parce que, à spécificité égale, le nombre de faux positifs reste élevé, tandis que le nombre de vrais positifs baisse avec la prévalence.

S4. UH = 1 × 1 064 - 1 024 = 40 UH. La longueur est de 64 × 0,78 = 49,9 mm. L'option Exclude on edges exclut du comptage les objets qui touchent le bord de l'image, car ils sont incomplets et fausseraient l'aire moyenne. Le nombre d'objets est comparé à la valeur obtenue par l'enseignant sur le même fichier.

S5. (1) HL7 v2, avec un code LOINC pour l'analyse. (2) DICOM, service C-STORE. (3) HL7 FHIR par l'interface REST, avec autorisation SMART App Launch. (4) Le résumé international du patient (IPS), via MyHealth@EU, obligatoire dans l'UE à partir du 26 mars 2029. (5) La signature électronique qualifiée, au sens d'eIDAS. Dans le JSON : le patient est la référence `Patient/demo-17`, le paramètre est le pouls (LOINC 8867-4), la valeur est 88, l'unité /min, et le moment est le 12 octobre 2026 à 8 h 30.

S6. Toute recommandation de prendre un antalgique et d'attendre, sans orientation immédiate vers les urgences, est classée comme signe d'alarme manqué, de gravité majeure. La référence correcte est : Walters WH, Wilder EI. Fabrication and errors in the bibliographic citations generated by ChatGPT. Sci Rep. 2023;13(1):14045. doi:10.1038/s41598-023-41032-5. La déclaration mentionne l'outil, la version, la date et l'objectif, et précise que les affirmations ont été vérifiées par l'auteur.

Tableau L14.1. Barème du sujet type (10 points). Modèle didactique.

Exercice	Ce qui est noté	Points	Ce qui annule les points
S1	dictionnaire 0,5 ; au moins trois validations 0,5 ; saut conditionnel 0,25 ; justification de l'absence d'e-mail 0,25	1,5	collecte du nom, du CNP ou de l'e-mail
S2	pour chaque cas : CIM-10 0,3 et CIM-11 0,3 ; versions notées 0,2	2,0	code d'extension seul ou en première position
S3	cinq indicateurs 1,0 ; recalcul à 2 % 0,6 ; explication 0,4	2,0	formules sans cellules de calcul visibles
S4	attributs de l'en-tête 0,4 ; UH 0,3 ; longueur 0,4 ; comptage et explication 0,4	1,5	mesure sur une image exportée en JPEG
S5	cinq standards à 0,2 chacun ; éléments du JSON 0,5	1,5	aucun
S6	affirmations classées 0,5 ; PMID vérifiés 0,3 ; Vancouver 0,4 ; déclaration 0,3	1,5	"confirmé" sans source identifiable

Des points partiels sont accordés pour un raisonnement correct comportant une erreur de calcul. Un code correct sans version reçoit la moitié des points, comme au TP 4.

L14.3 Erreurs fréquentes

Pièges fréquents

- En S1 : champs numériques laissés en texte libre, sans validation, ou saut conditionnel placé sur une question de type réponse courte, pour laquelle Google Forms ne le permet pas.
- En S2 : arrêt à la catégorie à trois caractères, "&" entre deux codes souches, absence de version.
- En S3 : division par le nombre total de patients au lieu du nombre de malades ou de sujets sains, confusion entre VPP et sensibilité.
- En S4 : modification de l'échelle avec Set Scale, clic sur le bouton Apply dans Window/Level, mesure sur une capture d'écran.
- En S5 : "HL7" sans version, confusion entre DICOM et PACS, IPS décrit comme un dossier complet.
- En S6 : affirmation jugée vraie parce qu'elle paraît plausible, référence acceptée sans recherche dans PubMed, liste alphabétique au lieu de l'ordre de citation.
- À tout moment : données réelles de personnes dans des captures d'écran ou des fichiers.

Toutes ces erreurs ont été discutées dans les chapitres et les TP dont proviennent les exercices. La liste de compétences du tableau 14.1 peut servir d'autoévaluation avant l'examen.

L14.4 Plan de préparation

La révision se fait à partir des TP, et non des notes de cours. Le plan ci-dessous s'étend sur trois semaines.

1. **Trois semaines avant.** Refaites les TP 3 et 4 sur de nouveaux cas. Pour le TP 4, choisissez cinq diagnostics du chapitre 4 et codez-les sans regarder les corrigés, puis comparez.
2. **Deux semaines avant.** Refaites les calculs du TP 13 avec d'autres valeurs et d'autres prévalences, jusqu'à ce que le calcul d'une matrice prenne moins de cinq minutes. Refaites le TP 7 sur une image DICOM anonymisée et enregistrez la macro.
3. **Une semaine avant.** Reprenez le TP 8 et le chapitre 8 pour les standards. Construisez sur papier un tableau "situation, standard" d'au moins dix lignes. Refaites la vérification d'une

réponse générée, selon le TP 10, et rédigez trois références Vancouver sans gestionnaire de références.

4. **Les derniers jours.** Traitez le sujet type en temps limité, en 90 minutes. Corrigez-vous vous-mêmes à l'aide du tableau L14.1 et notez l'exercice sur lequel vous avez perdu le plus de points.
5. **Le jour de l'examen.** Vérifiez l'accès au compte institutionnel et vos mots de passe avant l'épreuve, sans les noter sur la feuille.

Les TP avec feuilles de calcul (TP 3, TP 12, TP 13) prolongent les protocoles de [6], repris et mis à jour pour Google Sheets/Excel 2026. Les refaire sur de nouvelles données constitue la révision la plus efficace, car l'examen change les valeurs, non la méthode.

Travail à rendre

À l'examen, on remet un seul document contenant les réponses aux exercices S1-S6, accompagné de la feuille de calcul (S1, S3), du lien vers le formulaire partagé avec l'enseignant en tant que Lecteur (S1) et des captures d'écran de Fiji sans données personnelles (S4). À la séance de préparation, on remet le sujet type traité et autocorrigé à l'aide du tableau L14.1, avec les points obtenus à chaque exercice. Barème indicatif : tableau L14.1, total 10 points.

Variante hors ligne

Sans réseau, les exercices se traitent comme suit. S1 se fait dans une feuille de calcul locale, avec des règles de validation des données et une colonne qui signale les réponses incompatibles, comme dans la variante hors ligne du TP 3. Pour S2, l'enseignant distribue des extraits imprimés de la liste tabulaire de la CIM-10 et les pages CIM-11 des entités utilisées, y compris la section de postcoordination. S3 se résout avec la calculatrice de poche. Pour S4, Fiji installé localement ouvre le fichier DICOM copié au préalable sur l'ordinateur, et l'image Blobs est remplacée par l'image synthétique du TP 7. S5 ne nécessite pas de réseau. Pour S6, les notices PubMed des références sont imprimées et mélangées à des références inexistantes, et l'étudiant les compare à la liste reçue.

Annexe A. Réponses aux questions d'autoévaluation

QCS désigne les questions à choix simple (une seule réponse correcte) et QCM les questions à choix multiples (deux à quatre réponses correctes).

Tableau A.1. Grille de réponses par chapitre.

Chapitre	Réponses correctes
1	1-c; 2-b; 3-c; 4-a,c,d; 5-a,b,c,e
2	1-d; 2-d; 3-b; 4-a,b,d; 5-a,b,c
3	1-b; 2-a; 3-b; 4-a,b,d; 5-a,b,c,d
4	1-d; 2-c; 3-b; 4-a,b,d; 5-a,b,d
5	1-b; 2-b; 3-c; 4-a,b,d; 5-a,d,e
6	1-d; 2-c; 3-b; 4-a,b,d; 5-a,b,d
7	1-d; 2-e; 3-b; 4-a,b,d,e; 5-a,b,c
8	1-c; 2-c; 3-d; 4-a,b,c,e; 5-a,b,d
9	1-b; 2-c; 3-b; 4-a,b,c,e; 5-a,b,d
10	1-c; 2-c; 3-c; 4-a,b,d; 5-a,c,d
11	1-e; 2-b; 3-b; 4-a,b,c; 5-a,c,d
12	1-c; 2-c; 3-d; 4-a,c,e; 5-a,b,c,e
13	1-c; 2-b; 3-d; 4-a,b,d; 5-b,c,d
14	1-d; 2-b; 3-b; 4-a,b,d,e; 5-a,b,d; 6-c; 7-b; 8-a,c,e; 9-b; 10-a,b,c; 11-a,b,c; 12-c; 13-a; 14-a,b,d; 15-c; 16-a,b,c,e; 17-b; 18-a,b,d; 19-b; 20-a,b,c; 21-c; 22-a,b,c; 23-a; 24-a; 25-a,b,d

Annexe B. Glossaire

- **Affection principale** (angl. *main condition*) : affection diagnostiquée à la fin de l'épisode de soins et principalement responsable du besoin de traitement ou d'examens. (chap. 4)
- **Aire sous la courbe** (angl. *area under the curve*, AUC) : probabilité qu'un malade choisi au hasard ait un score plus élevé qu'un sujet sain. Elle mesure la discrimination, non la calibration. (chap. 13)
- **Analyse de la courbe de décision** (angl. *decision curve analysis*) : méthode qui calcule le bénéfice net d'un modèle pour chaque probabilité seuil à partir de laquelle le clinicien interviendrait. (chap. 13)
- **Anonymisation** (angl. *anonymisation*) : suppression irréversible du lien entre les données et la personne. Les données anonymes sortent du champ du RGPD, mais une anonymisation réelle est difficile à obtenir. (chap. 5)
- **Application de bien-être** (angl. *wellness application*) : logiciel destiné à une personne physique pour le traitement de données de santé à des fins autres que les soins médicaux, défini dans l'EHDS. (chap. 10)
- **Apprentissage automatique** (angl. *machine learning*) : domaine de l'IA dans lequel les systèmes apprennent à résoudre des problèmes en analysant des données qui s'y rapportent. (chap. 13)
- **Archétype** (angl. *archetype*) : modèle réutilisable d'un concept clinique, par exemple la tension artérielle, défini par des cliniciens dans openEHR. (chap. 8)
- **Authentification multifacteur** (angl. *multi-factor authentication*) : authentification qui combine au moins deux facteurs, par exemple le mot de passe et un code généré sur le téléphone. (chap. 5, 11)
- **Biais algorithmique** (angl. *algorithmic bias*) : erreur systématique d'un algorithme dont les conséquences retombent de manière disproportionnée sur un groupe particulier de patients. (chap. 5, 13)
- **Biais d'automatisation** (angl. *automation bias*) : tendance à accepter la suggestion d'un système automatisé même lorsqu'elle contredit d'autres éléments de preuve. (chap. 7, 13)
- **Bit** (angl. *binary digit*) : quantité d'information obtenue lors du choix entre deux possibilités également probables. Huit bits forment un octet. (chap. 2)
- **Calibration** (angl. *calibration*) : concordance entre les probabilités affichées par un modèle et la fréquence réelle de l'événement. (chap. 13)
- **Cause initiale du décès** (angl. *underlying cause of death*) : maladie ou traumatisme qui a déclenché l'enchaînement des événements morbides ayant conduit au décès, ou circonstances de l'accident ou de la violence qui ont entraîné le traumatisme mortel. (chap. 4)
- **Chiffrement asymétrique** (angl. *public-key cryptography*) : chiffrement au moyen d'une paire de clés liées mathématiquement, l'une publique et l'autre privée, conservée uniquement par son titulaire. (chap. 11)
- **Chiffrement de bout en bout** (angl. *end-to-end encryption*) : chiffrement dans lequel le message n'est déchiffré que sur l'appareil du destinataire, sans que le serveur intermédiaire puisse le lire. (chap. 9, 11)
- **Chiffrement symétrique** (angl. *symmetric encryption*) : chiffrement dans lequel l'expéditeur et le destinataire utilisent la même clé secrète, par exemple AES. (chap. 11)

- **Classification** (angl. *classification*) : ensemble de classes exhaustives et mutuellement exclusives, qui regroupent des concepts qu'il n'est plus nécessaire de distinguer dans l'analyse, comme la CIM. (chap. 4)
- **Clé étrangère** (angl. *foreign key*) : colonne qui doit désigner une ligne existante d'une autre table. (chap. 3)
- **Clé primaire** (angl. *primary key*) : attribut ou groupe d'attributs qui identifie de façon unique chaque ligne d'une table. (chap. 3)
- **Cluster** (angl. *cluster*) : dans la CIM-11, groupe de codes liés par postcoordination qui décrit un seul concept clinique. (chap. 4)
- **Code d'extension** (angl. *extension code*) : code CIM-11 du chapitre X, qui ajoute des détails à un code souche et ne peut figurer ni seul ni en première position d'un cluster. (chap. 4)
- **Code souche** (angl. *stem code*) : code CIM-11 qui peut être utilisé seul. (chap. 4)
- **Complaisance** (angl. *sycophancy*) : tendance d'un modèle de langage à accepter la prémisse erronée de l'utilisateur. (chap. 10)
- **Confidentialité** (angl. *confidentiality*) : propriété de l'information de n'être accessible qu'aux personnes autorisées. Elle concerne l'information, à la différence de la vie privée, qui concerne la personne. (chap. 5)
- **Connaissance** (angl. *knowledge*) : résultat de l'analyse formelle ou informelle de l'information, qui englobe les résultats des études, les heuristiques et les modèles. (chap. 1, 2)
- **Consentement dynamique** (angl. *dynamic consent*) : consentement géré sur une plateforme numérique, par laquelle le participant met à jour ses préférences au fil du temps et peut retirer son accord. (chap. 5)
- **Consentement large** (angl. *broad consent*) : consentement qui couvre une catégorie de recherches futures, et non une seule étude. (chap. 5)
- **Dérive du jeu de données** (angl. *dataset shift*) : situation dans laquelle les données rencontrées en exploitation ne ressemblent plus à celles qui ont servi au développement d'un modèle, en raison de la technologie, de la population ou de la pratique. (chap. 13)
- **Dictionnaire de données** (angl. *data dictionary*, *codebook*) : description de chaque variable d'un jeu de données (nom, étiquette, type, unité, valeurs autorisées et code des valeurs manquantes). (chap. 3)
- **Dispositif portable** (angl. *wearable*) : capteur porté sur le corps, généralement au poignet, au doigt, sur la poitrine ou sur le bras, qui transmet les données à un téléphone. (chap. 10)
- **Donnée clinique** (angl. *datum*) : observation unique relative à un patient, décrite par le patient, le paramètre, la valeur, le moment et la méthode. (chap. 1, 2)
- **Données concernant la santé** (angl. *data concerning health*) : données à caractère personnel relatives à la santé physique ou mentale d'une personne, y compris la prestation de services de soins de santé, qui révèlent des informations sur son état de santé. Elles constituent une catégorie particulière de données au sens du RGPD. (chap. 5)
- **Données générées par le patient** (angl. *patient-generated health data*) : données cliniquement pertinentes, recueillies par le patient ou par ses appareils en dehors du cadre traditionnel des soins. (chap. 2, 10)
- **Dossier de santé électronique** (DSE ; angl. *electronic health record*, EHR) : dépôt électronique d'informations sur la santé et les soins d'une personne, doté de fonctions qui soutiennent l'acte médical et les autres utilisations légitimes des données. (chap. 8)

- **Dossier médical électronique** (angl. *electronic medical record*, EMR) : documentation électronique d'un seul établissement, cabinet ou hôpital. (chap. 8)
- **Dossier personnel de santé** (angl. *personal health record*, PHR) : dossier contrôlé par le patient, dans lequel celui-ci consulte ses résultats, communique avec l'équipe soignante et peut ajouter des données mesurées à domicile. (chap. 8)
- **Échantillonnage** (angl. *sampling*) : mesure de la valeur d'un signal analogique à intervalles de temps égaux. (chap. 6)
- **Entrepôt de données** (angl. *data warehouse*) : base dans laquelle les données des systèmes de routine sont copiées périodiquement, nettoyées, intégrées et organisées pour l'analyse. (chap. 3)
- **Entropie** (angl. *entropy*) : information moyenne attendue d'une variable, c'est-à-dire son incertitude avant l'observation, $H = -\Sigma\, p \cdot \log_2 p$. (chap. 2)
- **Exactitude** (angl. *accuracy*) : proportion des décisions correctes d'un test, (VP + VN) / total. Elle peut induire en erreur lorsque la maladie est rare. (chap. 13)
- **Fatigue face aux alertes** (angl. *alert fatigue*) : fait d'ignorer les alertes en raison du grand nombre d'avertissements non pertinents. (chap. 8, 12)
- **Fenêtrage** (angl. *windowing*) : choix du niveau et de la largeur de l'intervalle d'unités Hounsfield affiché sur toute l'échelle de gris. Il ne modifie que l'affichage, non les données. (chap. 7)
- **Fonction de hachage** (angl. *hash function*) : fonction qui calcule, à partir de données de longueur quelconque, une empreinte de longueur fixe. Elle n'a pas de clé et sert à vérifier l'intégrité. (chap. 11)
- **Fonctionnement en mode dégradé** (angl. *downtime procedures*) : procédures selon lesquelles un hôpital fonctionne lorsque le système informatique est indisponible. (chap. 12)
- **Fondation de la CIM-11** (angl. *Foundation Component*) : base de connaissances sémantique dont sont extraites les linéarisations de la CIM-11. (chap. 4)
- **Fuite de données** (angl. *data leakage*) : accès du modèle, pendant son développement, à des informations dont il ne disposerait pas en utilisation réelle, ce qui gonfle artificiellement sa performance. (chap. 13)
- **Grand modèle de langage** (angl. *large language model*, LLM) : réseau de neurones entraîné sur de très grands volumes de texte pour prédire le fragment de mot suivant. Il ne consulte pas une base de faits vérifiés. (chap. 10, 13)
- **Hallucination** (angl. *hallucination*) : affirmation présentée comme un fait par un modèle génératif, mais non étayée par la source ou par la réalité. (chap. 10, 13)
- **Hameçonnage** (angl. *phishing*) : message qui imite une source de confiance pour obtenir un mot de passe, un clic ou l'ouverture d'une pièce jointe. (chap. 5, 11)
- **Horodatage** (angl. *time stamp*) : données électroniques qui associent un ensemble de données à un instant particulier et prouvent que ces données existaient à cet instant. (chap. 11)
- **IA explicable** (angl. *explainable AI*, XAI) : méthodes qui tentent de montrer ce qui a déterminé la prédiction d'un modèle, par exemple les valeurs SHAP ou les cartes de saillance. (chap. 13)
- **Indice de case-mix** (angl. *case-mix index*) : moyenne des poids relatifs des cas sortis d'un hôpital au cours d'une période. (chap. 4, 12)
- **Information** (angl. *information*) : données analysées et organisées de façon à prendre sens. (chap. 1, 2)

- **Informatique biomédicale** (angl. *biomedical informatics*) : domaine interdisciplinaire qui étudie et vise l'utilisation efficace des données, des informations et des connaissances biomédicales pour la recherche, la résolution de problèmes et la prise de décision, dans le but d'améliorer la santé. (chap. 1)
- **Informatique médicale** (angl. *medical informatics*) : application de l'informatique biomédicale centrée sur la maladie et sur l'activité du médecin. Dans cet ouvrage, la discipline telle que la rencontre l'étudiant en médecine. (chap. 1)
- **Interopérabilité** (angl. *interoperability*) : capacité de deux systèmes à échanger des données et à les utiliser sans intervention manuelle, aux niveaux technique, syntaxique, sémantique et organisationnel. (chap. 8)
- **Interopérabilité sémantique** (angl. *semantic interoperability*) : capacité du système qui reçoit les données à en comprendre le sens exact dans lequel elles ont été envoyées, grâce à des terminologies communes. (chap. 4, 8)
- **Journal d'audit** (ou piste d'audit ; angl. *audit trail*, *audit log*) : enregistrement de chaque accès aux données et de chaque modification, avec l'auteur, le moment et l'action. (chap. 3, 5, 11)
- **Lac de données** (angl. *data lake*) : dépôt qui conserve les données sous leur forme brute et ne les structure qu'au moment de l'analyse. (chap. 3)
- **Linéarisation** (angl. *linearization*) : extraction des entités de la Fondation de la CIM-11 dans une classification destinée à un usage précis, sans modification de leur sens. La MMS est la linéarisation statistique. (chap. 4)
- **Logiciel dispositif médical** (angl. *medical device software*) : logiciel ayant une finalité médicale déclarée par le fabricant pour un patient individuel, réglementé par le MDR. (chap. 10)
- **Maîtrise de l'IA** (angl. *AI literacy*) : connaissances et compétences du personnel qui utilise des systèmes d'IA, que les fournisseurs et les déployeurs doivent favoriser conformément à l'article 4 de l'AI Act. (chap. 13, 14)
- **Matrice de confusion** (angl. *confusion matrix*) : tableau à quatre cases (vrai positif, faux positif, faux négatif, vrai négatif) à partir duquel on calcule les indicateurs de performance d'un test. (chap. 13)
- **Modèle de fondation** (angl. *foundation model*) : modèle entraîné par auto-supervision sur de grands ensembles de données variés, capable d'accomplir de nombreuses tâches avec peu ou pas d'étiquettes spécifiques. (chap. 13)
- **Moindre privilège** (angl. *least privilege*) : attribution des accès par rôle, limitée aux besoins de ce rôle. (chap. 5, 11)
- **Moteur d'interfaces** (angl. *integration engine*) : serveur de communication situé au centre d'une architecture en étoile, qui reçoit, met en file d'attente, transforme et distribue les messages échangés entre les applications de l'hôpital. (chap. 12)
- **Nomenclature** (angl. *nomenclature*) : ensemble de termes codés qui nomment le plus exactement possible les concepts médicaux, sans les regrouper, comme SNOMED CT. (chap. 4)
- **Normalisation** (angl. *normalization*) : décomposition d'une table redondante en tables plus petites, liées par des clés, de sorte que chaque fait ne soit écrit qu'une seule fois. (chap. 3)

- **Ontologie** (angl. *ontology*) : représentation des concepts d'un domaine par des relations formelles ("est un", "a pour localisation"), à partir desquelles un programme peut déduire des relations qui ne sont pas écrites explicitement. (chap. 2, 4)
- **Pixel** (angl. *picture element*) : élément d'une image numérique bidimensionnelle, qui représente l'intensité d'une petite surface. L'élément de volume correspondant s'appelle voxel. (chap. 7)
- **Plagiat** (angl. *plagiarism*) : fait de présenter comme une création personnelle des textes, des idées, des données ou des méthodes empruntés aux œuvres d'autres auteurs, sans renvoi à la source. (chap. 5)
- **Postcoordination** (angl. *postcoordination*) : association de deux codes souches, ou d'un code souche et de codes d'extension, pour décrire un concept clinique, à l'aide de "&" et de "/" dans la CIM-11. (chap. 4)
- **Principes FAIR** (angl. *FAIR principles*) : exigence selon laquelle les données doivent être faciles à trouver, accessibles, interopérables et réutilisables. Cela ne signifie pas que les données sont publiques. (chap. 3)
- **Profil d'intégration IHE** (angl. *IHE integration profile*) : spécification qui précise quels standards et quelles options sont utilisées dans un scénario clinique donné, par exemple Scheduled Workflow. (chap. 7, 11)
- **Pseudonymisation** (angl. *pseudonymisation*) : traitement des données de telle façon qu'elles ne puissent plus être attribuées à une personne sans recours à des informations supplémentaires, conservées séparément. Les données pseudonymisées restent des données personnelles. (chap. 5)
- **Rapport signal/bruit** (angl. *signal-to-noise ratio*, SNR) : rapport entre la puissance du signal utile et la puissance du bruit, exprimé en décibels. (chap. 6)
- **Redondance** (angl. *redundancy*) : répétition partielle de l'information, qui permet de détecter et parfois de corriger les erreurs produites par le bruit. (chap. 2)
- **Règle 11** (angl. *rule 11*) : règle de l'annexe VIII du MDR selon laquelle on détermine la classe de risque d'un logiciel dispositif médical. (chap. 10, 13)
- **Repliement spectral** (angl. *aliasing*) : apparition d'une fausse fréquence, plus basse, lorsqu'un signal est échantillonné à moins du double de sa fréquence maximale. Il ne peut plus être corrigé après la numérisation. (chap. 6)
- **Ressource FHIR** (angl. *FHIR resource*) : unité de base de l'échange dans FHIR, comportant des éléments, des types de données et une section lisible par l'humain, représentée en JSON ou en XML. (chap. 8)
- **Résumé international du patient** (angl. *International Patient Summary*, IPS) : ensemble minimal de données cliniques destiné aux soins au-delà des frontières organisationnelles, avec trois sections obligatoires : problèmes, allergies, médicaments. (chap. 8)
- **Sensibilité** (angl. *sensitivity*) : proportion des malades détectés par le test, VP / (VP + FN). (chap. 13)
- **Signature électronique qualifiée** (angl. *qualified electronic signature*) : signature avancée créée à l'aide d'un dispositif qualifié et reposant sur un certificat qualifié. Elle a l'effet juridique d'une signature manuscrite. (chap. 11)
- **Spécificité** (angl. *specificity*) : proportion des sujets sains reconnus comme tels par le test, VN / (VN + FP). (chap. 13)

- **Système d'information** (angl. *information system*) : ensemble du traitement des données, de l'information et des connaissances dans une unité, avec les acteurs humains ou techniques qui le réalisent. La partie assistée par ordinateur constitue le système informatique. (chap. 11, 12)
- **Système d'information hospitalier** (SIH ; angl. *hospital information system*, HIS) : partie informatisée du système d'information de l'hôpital, composée de modules tels que ADT, FOCG, CPOE, LIS et RIS/PACS. (chap. 12)
- **Taux d'occupation des lits** (angl. *bed occupancy rate*) : journées d'hospitalisation réalisées divisées par les journées-lits disponibles, le tout multiplié par 100. (chap. 12)
- **Téléassistance médicale** (angl. *medical teleassistance*) : dans les normes roumaines, intégration en temps réel de données ou d'avis de professionnels situés à distance pour la réalisation d'actes médicaux. (chap. 9)
- **Téléconsultation** (angl. *teleconsultation*) : consultation à distance entre le patient et le médecin, généralement synchrone, par vidéo ou par téléphone. (chap. 9)
- **Téléexpertise** (angl. *tele-expertise*) : consultation à distance entre professionnels, synchrone ou asynchrone, pour obtenir un second avis. (chap. 9)
- **Télémédecine** (angl. *telemedicine*) : fourniture de services de santé à distance, au moyen des technologies de l'information et de la communication, à des fins de diagnostic, de traitement, de surveillance ou de prévention. (chap. 9)
- **Télésurveillance** (angl. *remote monitoring*) : transmission des paramètres mesurés par des dispositifs médicaux à une équipe qui reçoit des alertes lorsque les valeurs sortent des intervalles définis. (chap. 9)
- **Terminologie clinique** (angl. *clinical terminology*) : nomenclature contrôlée, avec des identifiants stables, des synonymes et des relations entre concepts. (chap. 4)
- **Triade CIA** (angl. *CIA triad*) : confidentialité, intégrité et disponibilité, les trois propriétés par lesquelles on décrit la sécurité de l'information. (chap. 5, 11)
- **Unité Hounsfield** (angl. *Hounsfield unit*) : unité des valeurs de tomodensitométrie, l'eau étant à 0 UH et l'air à environ -1 000 UH. (chap. 7)
- **Utilisation primaire** (angl. *primary use*) : utilisation des données de santé pour les soins de la personne concernée, y compris dans un autre État membre. (chap. 5, 8)
- **Utilisation secondaire** (angl. *secondary use*) : réutilisation des données de santé pour la recherche, la santé publique, la statistique, l'élaboration des politiques ou l'enseignement, par l'intermédiaire des organismes responsables de l'accès aux données. (chap. 5, 8)
- **Valeur prédictive positive** (angl. *positive predictive value*) : probabilité qu'un patient dont le test est positif soit malade, VP / (VP + FP). Elle dépend de la prévalence. (chap. 13)
- **Validation externe** (angl. *external validation*) : évaluation d'un modèle sur des données recueillies dans un autre lieu ou à une autre période que celles utilisées pour son développement. (chap. 7, 13)
- **Variabilité de la fréquence cardiaque** (VFC ; angl. *heart rate variability*, HRV) : fluctuation des intervalles entre les battements sinusaux normaux, décrite par des indicateurs tels que SDNN et RMSSD. (chap. 6)

Références bibliographiques

1. Mihalaş GI, Lungeanu D. Informatică medicală și biostatistică. Timişoara: Editura Victor Babeș; 2009.
2. Focșa MA. Informatica medicală în era digitală: de la date la inteligența artificială. Timișoara: Editura Victor Babeș; 2025.
3. Mihalaş GI. Inteligența artificială în medicină: ghid de conversație pentru medici, studenți și pacienți. Timișoara: Editura Victor Babeș; 2025.
4. Focșa M, Tudor A, Câtu C, Muntean C, Mada L. Informatică medicală: îndrumător de lucrări practice. Timișoara: Editura Victor Babeș; 2010.
5. Vernic C, Mada L, Lungeanu D, Muntean C, Apostol AS, Câtu CO, et al. Aplicații practice de informatică medicală și biostatistică. Timișoara: Editura Victor Babeș; 2010.
6. Vernic CV, Frandeș M, Tudor A, Apostol SA, Voicu A, Șerban C, et al. Prelucrarea datelor biomedicale: aplicații practice. Timișoara: Eurostampa; 2018.
7. Shortliffe EH, Chiang MF. Biomedical Informatics: The Science and the Pragmatics. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 3-44. doi:10.1007/978-3-030-58721-5_1
8. Mooney SD, Tenenbaum JD, Altman RB. Bioinformatics. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 273-98. doi:10.1007/978-3-030-58721-5_9
9. LaVenture M, Ross DA, Staes C, Yasnoff WA. Population and Public Health Informatics. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 613-36. doi:10.1007/978-3-030-58721-5_18
10. European Federation for Medical Informatics. The Romanian Society of Medical Informatics [Internet]. EFMI; 2021 Sep 20 [cited 2026 Sep 27]. Available from: https://efmi.org/2021/09/20/the-romanian-society-of-medical-informatics/
11. Shortliffe EH, Chiang MF. Biomedical Data: Their Acquisition, Storage, and Use. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 45-75. doi:10.1007/978-3-030-58721-5_2
12. Grossi E. The long journey of artificial intelligence in medicine: an overview. Clin Exp Rheumatol. 2025;43(5):815-21. doi:10.55563/clinexprheumatol/oamfed
13. Hersh W. Information Retrieval. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 755-94. doi:10.1007/978-3-030-58721-5_23
14. Musen MA, Middleton B, Greenes RA. Clinical Decision-Support Systems. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 795-840. doi:10.1007/978-3-030-58721-5_24
15. Jaffe C, Nguyen V, Kubick WR, Cooper T, Leftwich RB, Hammond WE. Standards in Biomedical Informatics. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 205-40. doi:10.1007/978-3-030-58721-5_7
16. Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. Gastrointest Endosc. 2020;92(4):807-12. doi:10.1016/j.gie.2020.06.040
17. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-9. doi:10.1038/s41586-021-03819-2
18. Mugaanyi J, Cai L, Cheng S, Lu C, Huang J. Evaluation of Large Language Model Performance and Reliability for Citations and References in Scholarly Writing: Cross-Disciplinary Study. J Med Internet Res. 2024;26:e52935. doi:10.2196/52935
19. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLOS Digit Health. 2023;2(2):e0000198. doi:10.1371/journal.pdig.0000198
20. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, et al. Large language models encode clinical knowledge. Nature. 2023;620(7972):172-80. doi:10.1038/s41586-023-06291-2
21. Peek N, Sujan M, Scott P. Digital health and care: emerging from pandemic times. BMJ Health Care Inform. 2023;30(1):e100861. doi:10.1136/bmjhci-2023-100861
22. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No

167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). Official Journal of the European Union. 2024 Jul 12;L 2024/1689.

23. Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847. Official Journal of the European Union. 2025 Mar 5;L 2025/327.
24. Rubin DL, Greenspan H, Hoogi A. Biomedical Imaging Informatics. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 299-362. doi:10.1007/978-3-030-58721-5_10
25. Sheikh A, Anderson M, Albala S, Casadei B, Franklin BD, Richards M, et al. Health information technology and digital innovation for national learning health and care systems. Lancet Digit Health. 2021;3(6):e383-96. doi:10.1016/S2589-7500(21)00005-4
26. Cronin RM, Jimison H, Johnson KB. Personal Health Informatics. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 363-89. doi:10.1007/978-3-030-58721-5_11
27. Tenenbaum JD, Shah NH, Altman RB. Translational Bioinformatics. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 867-911. doi:10.1007/978-3-030-58721-5_26
28. Gulkesen KH, Sonuvar ET. Data Privacy in Medical Informatics and Electronic Health Records: A Bibliometric Analysis. Health Care Anal. 2025. doi:10.1007/s10728-025-00519-0
29. Bichel-Findlay J, Koch S, Mantas J, Abdul SS, Al-Shorbaji N, Ammenwerth E, et al. Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics: Second Revision. Int J Med Inform. 2023;170:104908. doi:10.1016/j.ijmedinf.2022.104908
30. Car J, Ong QC, Erlikh Fox T, Leightley D, Kemp SJ, Švab I, et al. The Digital Health Competencies in Medical Education Framework: An International Consensus Statement Based on a Delphi Study. JAMA Netw Open. 2025;8(1):e2453131. doi:10.1001/jamanetworkopen.2024.53131
31. Khurana MP, Raaschou-Pedersen DE, Kurtzhals J, Bardram JE, Ostrowski SR, Bundgaard JS. Digital health competencies in medical school education: a scoping review and Delphi method study. BMC Med Educ. 2022;22(1):129. doi:10.1186/s12909-022-03163-7
32. Zhang W, Han J, Han X, Xu H, Wang L, Lin T, et al. Medical Students' Attitudes, Perceptions, and Self-Reported Familiarity With AI in Health Care: Systematic Review and Meta-Analysis. JMIR Med Educ. 2026;12:e89411. doi:10.2196/89411
33. Vuorikari R, Kluzer S, Punie Y. DigComp 2.2: The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes. EUR 31006 EN. Luxembourg: Publications Office of the European Union; 2022. doi:10.2760/115376
34. European Commission, Joint Research Centre. DigComp 3.0 [Internet]. Seville: JRC; 2025 [cited 2026 Sep 27]. Available from: https://joint-research-centre.ec.europa.eu/projects-and-activities/education-and-training/digital-transformation-education/digital-competence-framework-digcomp/digcomp-30_en
35. Williams GA, Falkenbach M, van Schaaijk A, Batenburg R, Wismar M; BeWell Consortium. Closing the digital skills gap in health care: identifying core digital skills and competencies and education and training opportunities for health professionals in the European Union. Policy Brief 71. Copenhagen: European Observatory on Health Systems and Policies, WHO Regional Office for Europe; 2025. Available from: https://www.ncbi.nlm.nih.gov/books/NBK618031/
36. World Health Organization. Global strategy on digital health 2020-2025. Geneva: WHO; 2021. ISBN 978-92-4-002092-4. Available from: https://www.who.int/publications/i/item/9789240020924
37. World Health Organization. World Health Assembly endorses extension of the global digital health strategy to 2027 [Internet]. Geneva: WHO; 2025 May 23 [cited 2026 Sep 27]. Available from: https://www.who.int/news/item/23-05-2025-world-health-assembly-endorses-extension-of-the-global-digital-health-strategy-to-2027
38. World Health Assembly. Decision WHA78(22). Global strategy on digital health 2020-2025: extension [Internet]. Geneva: World Health Organization; 2025 May 27 [cited 2026 Sep 27]. Available from: https://apps.who.int/gb/ebwha/pdf_files/WHA78/A78_(22)-en.pdf
39. National Library of Medicine (US). PubMed [Internet]. Bethesda (MD): National Center for Biotechnology Information; 2026 [cited 2026 Sep 27]. Available from: https://pubmed.ncbi.nlm.nih.gov/
40. Giustini D, Amar-Zifkin A, Chen E, Kung JY. Automated indexing in MEDLINE and the Medical Text Indexer (MTI), 2000-2025: a scoping review. J Med Libr Assoc. 2026;114(3):191-207. doi:10.5195/jmla.2026.2406

41. National Library of Medicine (US). PubMed User Guide [Internet]. Bethesda (MD): NLM; 2026 [cited 2026 Sep 27]. Available from: https://pubmed.ncbi.nlm.nih.gov/help/
42. Wallace SS, Barak G, Truong G, Parker MW. Hierarchy of Evidence Within the Medical Literature. Hosp Pediatr. 2022;12(8):745-50. doi:10.1542/hpeds.2022-006690
43. Bhattacharyya M, Miller VM, Bhattacharyya D, Miller LE. High Rates of Fabricated and Inaccurate References in ChatGPT-Generated Medical Content. Cureus. 2023;15(5):e39238. doi:10.7759/cureus.39238
44. Chelli M, Descamps J, Lavoué V, Trojani C, Azar M, Deckert M, et al. Hallucination Rates and Reference Accuracy of ChatGPT and Bard for Systematic Reviews: Comparative Analysis. J Med Internet Res. 2024;26:e53164. doi:10.2196/53164
45. Asai A, He J, Shao R, Shi W, Singh A, Chang JC, et al. Synthesizing scientific literature with retrieval-augmented language models. Nature. 2026;650(8103):857-63. doi:10.1038/s41586-025-10072-4
46. Carl N, Hetz MJ, Wies C, Haggenmüller S, Winterstein JT, Mangold MH, et al. Enhancing clinicians' trust in large language models via transparent source attribution: A randomized controlled evaluation in uro-oncology. Eur J Cancer. 2026;233:116168. doi:10.1016/j.ejca.2025.116168
47. Bedi S, Liu Y, Orr-Ewing L, Dash D, Koyejo S, Callahan A, et al. Testing and Evaluation of Health Care Applications of Large Language Models: A Systematic Review. JAMA. 2025;333(4):319-28. doi:10.1001/jama.2024.21700
48. International Medical Informatics Association. Romanian Society of Medical Informatics [Internet]. IMIA; [cited 2026 Sep 27]. Available from: https://imia-medinfo.org/wp/romanian-society-of-medical-informatics/
49. Societatea Română de Informatică Medicală. Romanian Society of Medical Informatics (SRIM) [Internet]. România: SRIM; 2026 [cited 2026 Sep 27]. Available from: https://www.srimed.ro/
50. Stoicu-Tivadar L, Gallos P, Bolboacă SD. Editorial. Appl Med Inform. 2024;46(Suppl 2):S1. Available from: https://ami.info.umfcluj.ro/index.php/AMI/article/view/1071
51. Applied Medical Informatics [Internet]. Cluj-Napoca: Universitatea de Medicină şi Farmacie "Iuliu Haţieganu", Department of Medical Informatics and Biostatistics; 2026 [cited 2026 Sep 27]. Available from: https://ami.info.umfcluj.ro/index.php/AMI
52. Asociația de Standardizare din România. ASRO/CT 319 - Informatică medicală [Internet]. Bucureşti: ASRO; 2026 [cited 2026 Sep 27]. Available from: https://www.asro.ro/asro-ct-319-informatica-medicala/
53. Hayes CN, Nakahara H, Ono A, Tsuge M, Oka S. From Omics to Multi-Omics: A Review of Advantages and Tradeoffs. Genes (Basel). 2024;15(12):1551. doi:10.3390/genes15121551
54. Tiase VL, Hull W, McFarland MM, Sward KA, Del Fiol G, Staes C, et al. Patient-generated health data and electronic health record integration: a scoping review. JAMIA Open. 2020;3(4):619-27. doi:10.1093/jamiaopen/ooaa052
55. Lewis AE, Weiskopf N, Abrams ZB, Foraker R, Lai AM, Payne PRO, et al. Electronic health record data quality assessment and tools: a systematic review. J Am Med Inform Assoc. 2023;30(10):1730-40. doi:10.1093/jamia/ocad120
56. Syed R, Eden R, Makasi T, Chukwudi I, Mamudu A, Kamalpour M, et al. Digital Health Data Quality Issues: Systematic Review. J Med Internet Res. 2023;25:e42615. doi:10.2196/42615
57. Benish WA. A Review of the Application of Information Theory to Clinical Diagnostic Testing. Entropy (Basel). 2020;22(1):97. doi:10.3390/e22010097
58. Owens DK, Goldhaber-Fiebert JD, Sox HC. Biomedical Decision Making: Probabilistic Clinical Reasoning. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 77-120. doi:10.1007/978-3-030-58721-5_3
59. Deane KD, Van Hoovels L, Joy VE, Olschowka N, Bossuyt X. From autoantibody test results to decision making: Incorporating likelihood ratios within medical practice. Autoimmun Rev. 2024;23(5):103537. doi:10.1016/j.autrev.2024.103537
60. Tsalatsanis A, Hozo I, Djulbegovic B. Research synthesis of information theory measures of uncertainty: Meta-analysis of entropy and mutual information of diagnostic tests. J Eval Clin Pract. 2021;27(2):246-55. doi:10.1111/jep.13475
61. Rohlfsen C, Shannon K, Parsons AS. Entropy in Clinical Decision-Making: A Narrative Review Through the Lens of Decision Theory. J Gen Intern Med. 2025;40(16):4033-9. doi:10.1007/s11606-025-09868-x
62. Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. doi:10.1007/978-3-030-58721-5

63. The Unicode Consortium. The Unicode Standard, Version 18.0.0 [Internet]. South San Francisco (CA): The Unicode Consortium; 2026 [cited 2026 Sep 27]. Available from: https://www.unicode.org/versions/Unicode18.0.0/
64. The Unicode Consortium. FAQ: UTF-8, UTF-16, UTF-32 and BOM [Internet]. South San Francisco (CA): The Unicode Consortium; c2026 [cited 2026 Sep 27]. Available from: https://www.unicode.org/faq/utf_bom.html
65. The Unicode Consortium. Latin Extended-B, Range: 0180-024F. The Unicode Standard, Version 18.0 [Internet]. South San Francisco (CA): The Unicode Consortium; 2026 [cited 2026 Sep 27]. Available from: https://www.unicode.org/charts/PDF/U0180.pdf
66. Guardiolle V, Bazoge A, Morin E, Daille B, Toublant D, Bouzillé G, et al. Linking Biomedical Data Warehouse Records With the National Mortality Database in France: Large-scale Matching Algorithm. JMIR Med Inform. 2022;10(11):e36711. doi:10.2196/36711
67. Wagner P, Strodthoff N, Bousseljot RD, Kreiseler D, Lunze FI, Samek W, et al. PTB-XL, a large publicly available electrocardiography dataset. Sci Data. 2020;7(1):154. doi:10.1038/s41597-020-0495-6
68. Erickson BJ. Imaging Systems in Radiology. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 733-53. doi:10.1007/978-3-030-58721-5_22
69. Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, et al. The complete sequence of a human genome. Science. 2022;376(6588):44-53. doi:10.1126/science.abj6987
70. Betschart RO, Sandberg F, Blankenberg S, Zoche M, Zeller T, Ziegler A. A benchmark study of compression software for human short-read sequence data. Sci Rep. 2025;15(1):15358. doi:10.1038/s41598-025-00491-8
71. Choi BCK, Barengo NC, Diaz PA. Public health surveillance and the data, information, knowledge, intelligence and wisdom paradigm. Rev Panam Salud Publica. 2024;48:e9. doi:10.26633/RPSP.2024.9
72. Hänsel K, Dudgeon SN, Cheung KH, Durant TJS, Schulz WL. From Data to Wisdom: Biomedical Knowledge Graphs for Real-World Data Insights. J Med Syst. 2023;47(1):65. doi:10.1007/s10916-023-01951-2
73. Serrano-Zamago AB, Altamirano-Bustamante MM. Appealing to Tacit Knowledge and Axiology to Enhance Medical Practice in the COVID-19 Pandemic: A Systematic Review and Hermeneutic Bioethical Analysis. Front Public Health. 2021;9:686773. doi:10.3389/fpubh.2021.686773
74. Charlet J, Cui L, Section Editors for the IMIA Yearbook Section on Knowledge Representation and Management. Knowledge Representation and Management 2022: Findings in Ontology Development and Applications. Yearb Med Inform. 2023;32(1):225-9. doi:10.1055/s-0043-1768747
75. Iglesias N, Juarez JM, Campos M. Comprehensive analysis of rule formalisms to represent clinical guidelines: Selection criteria and case study on antibiotic clinical guidelines. Artif Intell Med. 2020;103:101741. doi:10.1016/j.artmed.2019.101741
76. Ahn JW, Jang HY, Son YJ. Critical care nurses' communication challenges during handovers: A systematic review and qualitative meta-synthesis. J Nurs Manag. 2021;29(4):623-34. doi:10.1111/jonm.13207
77. Troyer L, Brady W. Barriers to effective EMS to emergency department information transfer at patient handover: A systematic review. Am J Emerg Med. 2020;38(7):1494-503. doi:10.1016/j.ajem.2020.04.036
78. Humphrey KE, Sundberg M, Milliren CE, Graham DA, Landrigan CP. Frequency and Nature of Communication and Handoff Failures in Medical Malpractice Claims. J Patient Saf. 2022;18(2):130-7. doi:10.1097/PTS.0000000000000937
79. Starmer AJ, Spector ND, O'Toole JK, Bismilla Z, Calaman S, Campos ML, et al. Implementation of the I-PASS handoff program in diverse clinical environments: A multicenter prospective effectiveness implementation study. J Hosp Med. 2023;18(1):5-14. doi:10.1002/jhm.12979
80. McCarthy S, Motala A, Lawson E, Shekelle PG. Use of structured handoff protocols for within-hospital unit transitions: a systematic review from Making Healthcare Safer IV. BMJ Qual Saf. 2025;34(10):680-90. doi:10.1136/bmjqs-2024-018385
81. Lo L, Rotteau L, Shojania K. Can SBAR be implemented with high fidelity and does it improve communication between healthcare workers? A systematic review. BMJ Open. 2021;11(12):e055247. doi:10.1136/bmjopen-2021-055247
82. Delardes B, McLeod L, Chakraborty S, Bowles KA. What is the effect of electronic clinical handovers on patient outcomes? A systematic review. Health Informatics J. 2020;26(4):2422-34. doi:10.1177/1460458220905162

83. Guyatt G, Agoritsas T, Brignardello-Petersen R, Mustafa RA, Rylance J, Foroutan F, et al. Core GRADE 1: overview of the Core GRADE approach. BMJ. 2025;389:e081903. doi:10.1136/bmj-2024-081903
84. Nussbaumer-Streit B, Booth A, Garritty C, Hamel C, Munn Z, Tricco AC, et al. Overview of evidence synthesis types and modes. J Clin Epidemiol. 2025;187:111970. doi:10.1016/j.jclinepi.2025.111970
85. Simmonds M, Elliott JH, Synnot A, Turner T. Living Systematic Reviews. Methods Mol Biol. 2022;2345:121-34. doi:10.1007/978-1-0716-1566-9_7
86. Cheyne S, Fraile Navarro D, Hill K, McDonald S, Tunnicliffe D, White H, et al. Methods for living guidelines: early guidance based on practical experience. Paper 1: Introduction. J Clin Epidemiol. 2023;155:84-96. doi:10.1016/j.jclinepi.2022.12.024
87. Lichtner G, Alper BS, Jurth C, Spies C, Boeker M, Meerpohl JJ, et al. Representation of evidence-based clinical practice guideline recommendations on FHIR. J Biomed Inform. 2023;139:104305. doi:10.1016/j.jbi.2023.104305
88. Rim SH, Colbert J, Kobernik E, Richards TB, Miller JW, Coleman King S, et al. Building Computable Guidelines and Clinical Decision Support Tools for Cervical, Breast and Colorectal Cancer Screening and Risk Management: A CDC Initiative Update. J Womens Health (Larchmt). 2026:15409996261489070. doi:10.1177/15409996261489070
89. Adam GP, Davies MC, George J, Caputo E, Htun JM, Coppola EL, et al. Machine Learning Tools To (Semi-)Automate Evidence Synthesis: A Rapid Review and Evidence Map. Research White Paper. AHRQ Publication No. 25-EHC006. Rockville (MD): Agency for Healthcare Research and Quality; 2024 Dec. doi:10.23970/AHRQEPCWHITEPAPERMACHINE. Available from: https://www.ncbi.nlm.nih.gov/books/NBK614977/
90. Khraisha Q, Put S, Kappenberg J, Warraitch A, Hadfield K. Can large language models replace humans in systematic reviews? Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. Res Synth Methods. 2024;15(4):616-26. doi:10.1002/jrsm.1715
91. Buko T, Tuczko N, Ishikawa T. DNA Data Storage. BioTech (Basel). 2023;12(2):44. doi:10.3390/biotech12020044
92. Chow JCL. Quantum Computing in Medicine. Med Sci (Basel). 2024;12(4):67. doi:10.3390/medsci12040067
93. HL7 International. FHIR Release 5 (v5.0.0): Resource Formats [Internet]. Ann Arbor (MI): Health Level Seven International; 2023 [cited 2026 Sep 27]. Available from: https://hl7.org/fhir/formats.html
94. Dilling TJ. Artificial Intelligence Research: The Utility and Design of a Relational Database System. Adv Radiat Oncol. 2020;5(6):1280-5. doi:10.1016/j.adro.2020.06.027
95. The PostgreSQL Global Development Group. PostgreSQL 18 Documentation. Appendix M. Glossary [Internet]. The PostgreSQL Global Development Group; 2025 [cited 2026 Sep 27]. Available from: https://www.postgresql.org/docs/current/glossary.html
96. Gamal A, Barakat S, Rezk A. Standardized electronic health record data modeling and persistence: A comparative review. J Biomed Inform. 2021;114:103670. doi:10.1016/j.jbi.2020.103670
97. Griffier R, Mougin F, Jouhet V. Integrating Health Care Data in an Informatics for Integrating Biology & the Bedside (i2b2) Model Persisted Through Elasticsearch: Design, Implementation, and Evaluation in a French University Hospital. JMIR Med Inform. 2025;13:e65753. doi:10.2196/65753
98. Idowu EAA, Teo J, Salih S, Valverde J, Yeung JA. Streams, rivers and data lakes: an introduction to understanding modern electronic healthcare records. Clin Med (Lond). 2023;23(4):409. doi:10.7861/clinmed.2022-0325
99. Padoan A, Plebani M. Flowing through laboratory clinical data: the role of artificial intelligence and big data. Clin Chem Lab Med. 2022;60(12):1875-80. doi:10.1515/cclm-2022-0653
100. Li X, Wood AR, Yuan Y, Zhang M, Huang Y, Hawkes G, et al. Streamlining large-scale genomic data management: Insights from the UK Biobank whole-genome sequencing data. Cell Genom. 2025;5(12):101009. doi:10.1016/j.xgen.2025.101009
101. Reich C, Ostropolets A, Ryan P, Rijnbeek P, Schuemie M, Davydov A, et al. OHDSI Standardized Vocabularies-a large-scale centralized reference ontology for international data harmonization. J Am Med Inform Assoc. 2024;31(3):583-90. doi:10.1093/jamia/ocad247
102. Santos HP, Colaço Júnior M, Valentim R, Dos Santos JPQ, da Silva MEM, de Oliveira LAHG, et al. Scientific and technological mapping of big data architectures in healthcare. Front Public Health. 2026;14:1775715. doi:10.3389/fpubh.2026.1775715
103. Observational Health Data Sciences and Informatics (OHDSI). OMOP Common Data Model [Internet]. OHDSI; 2026 [cited 2026 Sep 27]. Available from: https://ohdsi.github.io/CommonDataModel/

104. Quiroz JC, Chard T, Sa Z, Ritchie A, Jorm L, Gallego B. Extract, transform, load framework for the conversion of health databases to OMOP. PLoS One. 2022;17(4):e0266911. doi:10.1371/journal.pone.0266911
105. Voss EA, Blacketer C, van Sandijk S, Moinat M, Kallfelz M, van Speybroeck M, et al. European Health Data & Evidence Network-learnings from building out a standardized international health data network. J Am Med Inform Assoc. 2023;31(1):209-19. doi:10.1093/jamia/ocad214
106. Hripcsak G, Schuemie MJ, Madigan D, Ryan PB, Suchard MA. Drawing Reproducible Conclusions from Observational Clinical Data with OHDSI. Yearb Med Inform. 2021;30(1):283-9. doi:10.1055/s-0041-1726481
107. Reinecke I, Zoch M, Reich C, Sedlmayr M, Bathelt F. The Usage of OHDSI OMOP - A Scoping Review. Stud Health Technol Inform. 2021;283:95-103. doi:10.3233/SHTI210546
108. i2b2 tranSMART Foundation. i2b2: Informatics for Integrating Biology and the Bedside [Internet]. i2b2 tranSMART Foundation; 2026 [cited 2026 Sep 27]. Available from: https://www.i2b2.org/
109. GO FAIR Foundation. The FAIR Guiding Principles [Internet]. GO FAIR Foundation; 2026 [cited 2026 Sep 27]. Available from: https://www.gofair.foundation/fair-principles
110. Sinaci AA, Núñez-Benjumea FJ, Gencturk M, Jauer ML, Deserno T, Chronaki C, et al. From Raw Data to FAIR Data: The FAIRification Workflow for Health Research. Methods Inf Med. 2020;59(S 01):e21-32. doi:10.1055/s-0040-1713684
111. Inau ET, Sack J, Waltemath D, Zeleke AA. Initiatives, Concepts, and Implementation Practices of the Findable, Accessible, Interoperable, and Reusable Data Principles in Health Data Stewardship: Scoping Review. J Med Internet Res. 2023;25:e45013. doi:10.2196/45013
112. Schmidt CO, Struckmann S, Enzenbach C, Reineke A, Stausberg J, Damerow S, et al. Facilitating harmonized data quality assessments. A data quality framework for observational health research data collections with software implementations in R. BMC Med Res Methodol. 2021;21(1):63. doi:10.1186/s12874-021-01252-7
113. Parums DV. Editorial: Registries and Population Databases in Clinical Research and Practice. Med Sci Monit. 2021;27:e933554. doi:10.12659/MSM.933554
114. Malenka DJ, Bhatt DL, Bradley SM, Shahian DM, Draoui J, Segawa CA, et al. The National Cardiovascular Data Registry Data Quality Program 2020: JACC State-of-the-Art Review. J Am Coll Cardiol. 2022;79(17):1704-12. doi:10.1016/j.jacc.2022.02.034
115. Feier CVI, Santoro RR, Faur AM, Muntean C, Olariu S. Assessing Changes in Colon Cancer Care during the COVID-19 Pandemic: A Four-Year Analysis at a Romanian University Hospital. J Clin Med. 2023;12(20):6558. doi:10.3390/jcm12206558
116. Feier CVI, Bardan R, Muntean C, Olariu A, Olariu S. Impact of the COVID-19 Pandemic on the Elective Surgery for Colorectal Cancer: Lessons to Be Learned. Medicina (Kaunas). 2022;58(10):1322. doi:10.3390/medicina58101322
117. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). ICH Harmonised Guideline. Guideline for Good Clinical Practice E6(R3). Final version adopted on 6 January 2025 [Internet]. Geneva: ICH; 2025 [cited 2026 Sep 27]. Available from: https://database.ich.org/sites/default/files/ICH_E6%28R3%29_Step4_FinalGuideline_2025_0106.pdf
118. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Official Journal of the European Union. 2016 May 4;L 119:1.
119. Vanderbilt University Libraries. REDCap at Vanderbilt University [Internet]. Nashville (TN): Vanderbilt University; 2026 [cited 2026 Sep 27]. Available from: https://www.library.vanderbilt.edu/redcap/
120. Garcia KKS, Abrahão AA. Research Development Using REDCap Software. Healthc Inform Res. 2021;27(4):341-9. doi:10.4258/hir.2021.27.4.341
121. Van Bulck L, Wampers M, Moons P. Research Electronic Data Capture (REDCap): tackling data collection, management, storage, and privacy challenges. Eur J Cardiovasc Nurs. 2022;21(1):85-91. doi:10.1093/eurjcn/zvab104
122. Maré IA, Kramer B, Hazelhurst S, Nhlapho MD, Zent R, Harris PA, et al. Electronic Data Capture System (REDCap) for Health Care Research and Training in a Resource-Constrained Environment: Technology Adoption Case Study. JMIR Med Inform. 2022;10(8):e33402. doi:10.2196/33402
123. Google. How to set rules for your form. Google Docs Editors Help [Internet]. Mountain View (CA): Google; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/docs/answer/3378864

124. Google. Show questions based on answers. Google Docs Editors Help [Internet]. Mountain View (CA): Google; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/docs/answer/141062
125. Google. Choose where to save form responses. Google Docs Editors Help [Internet]. Mountain View (CA): Google; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/docs/answer/2917686
126. Zimba O, Gasparyan AY. Designing, Conducting, and Reporting Survey Studies: A Primer for Researchers. J Korean Med Sci. 2023;38(48):e403. doi:10.3346/jkms.2023.38.e403
127. Centers for Disease Control and Prevention. Epi Info [Internet]. Atlanta (GA): CDC; 2026 [updated 2026 Jun 22; cited 2026 Sep 27]. Available from: https://www.cdc.gov/epiinfo/index.html
128. Google. View and manage form responses. Google Docs Editors Help [Internet]. Mountain View (CA): Google; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/docs/answer/139706
129. Harrison JE, Weber S, Jakob R, Chute CG. ICD-11: an international classification of diseases for the twenty-first century. BMC Med Inform Decis Mak. 2021;21(Suppl 6):206. doi:10.1186/s12911-021-01534-6
130. Riley M, Lee J, Richardson S, Gjorgioski S, Robinson K. The applications of Australian-coded ICD-10 and ICD-10-AM data in research: A scoping review of the literature. Health Inf Manag. 2024;53(1):41-50. doi:10.1177/18333583231198592
131. Winter A, Ammenwerth E, Haux R, Marschollek M, Steiner B, Jahn F. Health Information Systems: Technological and Management Perspectives. 3rd ed. Cham: Springer; 2023. Chapter 3, Technological Perspective: Architecture, Integration, and Standards; p. 51-152. doi:10.1007/978-3-031-12310-8_3
132. Priyadharshini U, Vijayan R. Enhancing healthcare through ontology: a systematic review of challenges and future directions. Front Artif Intell. 2026;9:1872024. doi:10.3389/frai.2026.1872024
133. World Health Organization. International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10), version 2019, and Volume 2 Instruction manual, 6th ed. [Internet]. Geneva: WHO; 2019 [cited 2026 Sep 27]. Available from: https://icd.who.int/browse10/2019/en
134. World Health Organization. ICD-11 Reference Guide [Internet]. Geneva: WHO; 2022 [cited 2026 Sep 27]. Available from: https://icdcdn.who.int/icd11referenceguide/en/html/index.html
135. World Health Organization. ICD-11 Implementation: frequently asked questions [Internet]. Geneva: WHO; 2026 [cited 2026 Sep 27]. Available from: https://www.who.int/standards/classifications/frequently-asked-questions/icd-11-implementation
136. Fung KW, Xu J, Bodenreider O. The new International Classification of Diseases 11th edition: a comparative analysis with ICD-10 and ICD-10-CM. J Am Med Inform Assoc. 2020;27(5):738-46. doi:10.1093/jamia/ocaa030
137. Eastwood CA, Southern DA, Khair S, Doktorchik C, Cullen D, Ghali WA, et al. Field testing a new ICD coding system: methods and early experiences with ICD-11 Beta Version 2018. BMC Res Notes. 2022;15(1):343. doi:10.1186/s13104-022-06238-2
138. Lee H, Kim S. Impact of the ICD-11 on the accuracy of clinical coding in Korea. Health Inf Manag. 2023;52(3):221-8. doi:10.1177/18333583221095147
139. Institutul Național de Management al Serviciilor de Sănătate. Grupuri înrudite de diagnostic - clasificarea RO DRG v.1.1, versiunea martie 2024. Manual de definiții, volumul 1 [Internet]. București: INMSS; 2024 [cited 2026 Sep 27]. Available from: https://drg.inmss.ro/DocDRG/download.php?fi=28
140. World Health Organization. International Classification of Diseases (ICD) [Internet]. Geneva: WHO; 2026 [cited 2026 Sep 27]. Available from: https://www.who.int/standards/classifications/classification-of-diseases
141. World Health Organization. ICD-11 for Mortality and Morbidity Statistics, release 2026-01 [Internet]. Geneva: WHO; 2026 [cited 2026 Sep 27]. Available from: https://icd.who.int/browse/latest-release/mms/en
142. World Health Organization. ICD-11 Fact Sheet [Internet]. Geneva: WHO; 2026 May [cited 2026 Sep 27]. Available from: https://icd.who.int/en/docs/ICD11_Fact_Sheet_2026.pdf
143. Fung KW, Xu J, McConnell-Lamptey S, Pickett D, Bodenreider O. A practical strategy to use the ICD-11 for morbidity coding in the United States without a clinical modification. J Am Med Inform Assoc. 2023;30(10):1614-21. doi:10.1093/jamia/ocad128
144. Doktorchik C, Southern DA, King JA, Quan H. Applying an ICD-10 to ICD-11 mapping tool to identify causes of death codes in an Alberta dataset. Popul Health Metr. 2024;22(1):38. doi:10.1186/s12963-024-00358-6

145. Lyall K, Roberge I, Yang W, Baker S, Denny K. ICD-11 in Canada: Leveraging crosswalks to evaluate adoption, impact and transition strategies. Health Inf Manag. 2025:18333583251392355. doi:10.1177/18333583251392355
146. World Health Organization. International Classification of Functioning, Disability and Health (ICF) [Internet]. Geneva: WHO; 2026 [cited 2026 Sep 27]. Available from: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health
147. World Health Organization. International Classification of Health Interventions (ICHI) [Internet]. Geneva: WHO; 2026 [cited 2026 Sep 27]. Available from: https://www.who.int/standards/classifications/international-classification-of-health-interventions
148. SNOMED International. What is SNOMED CT? [Internet]. London: SNOMED International; 2026 [cited 2026 Sep 27]. Available from: https://www.snomed.org/what-is-snomed-ct
149. SNOMED International. SNOMED CT Starter Guide [Internet]. London: SNOMED International; 2026 [cited 2026 Sep 27]. Available from: https://docs.snomed.org/snomed-ct-practical-guides/snomed-ct-starter-guide
150. Chang E, Mostafa J. Use of SNOMED CT, 2020-2025: literature review. J Am Med Inform Assoc. 2026;33(9):1798-810. doi:10.1093/jamia/ocag094
151. SNOMED International. Our Members [Internet]. London: SNOMED International; 2026 [cited 2026 Sep 27]. Available from: https://www.snomed.org/members
152. Regenstrief Institute. LOINC Users Guide: 2 Major Parts of a LOINC term [Internet]. Indianapolis (IN): Regenstrief Institute; 2026 Aug 17 [cited 2026 Sep 27]. Available from: https://loinc.org/kb/users-guide/major-parts-of-a-loinc-term/
153. Regenstrief Institute. LOINC 2345-7 Glucose [Mass/volume] in Serum or Plasma [Internet]. Indianapolis (IN): Regenstrief Institute; 2026 [cited 2026 Sep 27]. Available from: https://loinc.org/2345-7
154. WHO Collaborating Centre for Drug Statistics Methodology. ATC: structure and principles [Internet]. Oslo: Norwegian Institute of Public Health; 2026 [cited 2026 Sep 27]. Available from: https://atcddd.fhi.no/atc/structure_and_principles/
155. WHO Collaborating Centre for Drug Statistics Methodology. DDD: definition and general considerations [Internet]. Oslo: Norwegian Institute of Public Health; 2026 [cited 2026 Sep 27]. Available from: https://atcddd.fhi.no/ddd/definition_and_general_considera/
156. National Library of Medicine (US). Medical Subject Headings (MeSH) [Internet]. Bethesda (MD): NLM; 2026 [cited 2026 Sep 27]. Available from: https://www.nlm.nih.gov/mesh/meshhome.html
157. National Library of Medicine (US). Unified Medical Language System (UMLS) [Internet]. Bethesda (MD): NLM; 2026 [cited 2026 Sep 27]. Available from: https://www.nlm.nih.gov/research/umls/index.html
158. Pi S, Masterson J, Ma SP, Corbin CK, Milstein A, Chen JH. Using Case Mix Index within Diagnosis-Related Groups to Evaluate Variation in Hospitalization Costs at a Large Academic Medical Center. AMIA Annu Symp Proc. 2023;2023:1201-8.
159. Liu J, Capurro D, Nguyen A, Verspoor K. Early prediction of diagnostic-related groups and estimation of hospital cost by processing clinical notes. NPJ Digit Med. 2021;4(1):103. doi:10.1038/s41746-021-00474-9
160. Chen YJ, Zhang XY, Yan JQ, Xue-Tang, Qian MC, Ying XH. Impact of Diagnosis-Related Groups on Inpatient Quality of Health Care: A Systematic Review and Meta-Analysis. Inquiry. 2023;60:469580231167011. doi:10.1177/00469580231167011
161. Campbell AC, Tyler J, Shuttleworth RR, Pearce LA, Heerde JA, Borschmann R, et al. Accuracy of administrative data in ascertaining health conditions: a systematic review. JAMIA Open. 2026;9(3):ooag109. doi:10.1093/jamiaopen/ooag109
162. Gabel E, Gal J, Grogan T, Hofer I. A retrospective analysis using comorbidity detecting algorithmic software to determine the incidence of International Classification of Diseases (ICD) code omissions and appropriateness of Diagnosis-Related Group (DRG) code modifiers. BMC Med Inform Decis Mak. 2024;24(1):309. doi:10.1186/s12911-024-02724-8
163. Crespin D, Dworsky M, Levin J, Ruder T, Whaley CM. Upcoding Linked To Up To Two-Thirds Of Growth In Highest-Intensity Hospital Discharges In 5 States, 2011-19. Health Aff (Millwood). 2024;43(12):1619-27. doi:10.1377/hlthaff.2024.00596
164. Khalid AR, Ali H, Fathima K, Sonia Carole K, Kim HC. Recent Advances in AI for Automated ICD Coding: A Systematic Literature Review. J Med Syst. 2026;50(1):113. doi:10.1007/s10916-026-02429-7

165. Soroush A, Glicksberg BS, Zimlichman E, Barash Y, Freeman R, Charney AW, et al. Large Language Models Are Poor Medical Coders - Benchmarking of Medical Code Querying. NEJM AI. 2024;1(5). doi:10.1056/aidbp2300040
166. Naliyatthaliyazchayil P, Muthyala R, Gichoya JW, Purkayastha S. Evaluating the Reasoning Capabilities of Large Language Models for Medical Coding and Hospital Readmission Risk Stratification: Zero-Shot Prompting Approach. J Med Internet Res. 2025;27:e74142. doi:10.2196/74142
167. Galland-Decker C, Ursenbacher M, Nunes C, Bouche F, Blanchard M, Carra G, et al. Improving Meditron for Medical Coding Through Fine-Tuning: A Comparative Evaluation Against GPT-4. Stud Health Technol Inform. 2026;336:1072-3. doi:10.3233/SHTI260356
168. Chomutare T, Svenning TO, Hernández MÁT, Ngo PD, Budrionis A, Markljung K, et al. Artificial Intelligence to Improve Clinical Coding Practice in Scandinavia: Crossover Randomized Controlled Trial. J Med Internet Res. 2025;27:e71904. doi:10.2196/71904
169. Friedman Y, Smajdor A. Synthetic Medical Data vs. The Four Principles: Ethical Trade-Offs and Zombie Data. Health Care Anal. 2026. doi:10.1007/s10728-026-00596-9
170. Goodman KW, Miller RA. Ethics in Biomedical and Health Informatics: Users, Standards, and Outcomes. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 391-423. doi:10.1007/978-3-030-58721-5_12
171. Legea nr. 46/2003 a drepturilor pacientului. Monitorul Oficial al României, Partea I. 2003 Jan 29;(51).
172. Cohen E, Byrom B, Becher A, Jörntén-Karlsson M, Mackenzie AK. Comparative Effectiveness of eConsent: Systematic Review. J Med Internet Res. 2023;25:e43883. doi:10.2196/43883
173. Zenker S, Strech D, Ihrig K, Jahns R, Müller G, Schickhardt C, et al. Data protection-compliant broad consent for secondary use of health care data and human biosamples for (bio)medical research: Towards a new German national standard. J Biomed Inform. 2022;131:104096. doi:10.1016/j.jbi.2022.104096
174. Lee AR, Koo D, Kim IK, Lee E, Kim HH, Yoo S, et al. Identifying facilitators of and barriers to the adoption of dynamic consent in digital health ecosystems: a scoping review. BMC Med Ethics. 2023;24(1):107. doi:10.1186/s12910-023-00988-9
175. Phuyal S, Bhandari M, Correia Bezerra R, Bista R, Ferreira JC. Dynamic Consent for Secondary Use of Health Data: Challenges and Opportunities Under European Law. JMIR Med Inform. 2026;14:e93348. doi:10.2196/93348
176. Legea nr. 190/2018 privind măsuri de punere în aplicare a Regulamentului (UE) 2016/679 al Parlamentului European și al Consiliului din 27 aprilie 2016 privind protecția persoanelor fizice în ceea ce privește prelucrarea datelor cu caracter personal și privind libera circulație a acestor date și de abrogare a Directivei 95/46/CE (Regulamentul general privind protecția datelor). Monitorul Oficial al României, Partea I. 2018 Jul 26;(651).
177. Olatunji IE, Rauch J, Katzensteiner M, Khosla M. A Review of Anonymization for Healthcare Data. Big Data. 2024;12(6):538-55. doi:10.1089/big.2021.0169
178. Packhäuser K, Gündel S, Münster N, Syben C, Christlein V, Maier A. Deep learning-based patient re-identification is able to exploit the biometric nature of medical chest X-ray data. Sci Rep. 2022;12(1):14851. doi:10.1038/s41598-022-19045-3
179. Xia W, Liu Y, Wan Z, Vorobeychik Y, Kantacioglu M, Nyemba S, et al. Enabling realistic health data re-identification risk assessment through adversarial modeling. J Am Med Inform Assoc. 2021;28(4):744-52. doi:10.1093/jamia/ocaa327
180. European Union Agency for Cybersecurity (ENISA). ENISA Threat Landscape: Health Sector (January 2021 to March 2023) [Internet]. Athens: ENISA; 2023 Jul [cited 2026 Sep 27]. Available from: https://www.enisa.europa.eu/publications/health-threat-landscape
181. Jalali MS, Bruckes M, Westmattelmann D, Schewe G. Why Employees (Still) Click on Phishing Links: Investigation in Hospitals. J Med Internet Res. 2020;22(1):e16775. doi:10.2196/16775
182. Herrera Montano I, Góngora Alonso S, Sañudo García S, García Aranda JJ, Rodrígues JJPC, de la Torre Díez I. Security practices and insider threats in Spanish healthcare centers: a survey-based risk assessment. Int J Med Inform. 2026;205:106107. doi:10.1016/j.ijmedinf.2025.106107
183. Neprash HT, McGlave CC, Cross DA, Virnig BA, Puskarich MA, Huling JD, et al. Trends in Ransomware Attacks on US Hospitals, Clinics, and Other Health Care Delivery Organizations, 2016-2021. JAMA Health Forum. 2022;3(12):e224873. doi:10.1001/jamahealthforum.2022.4873
184. Dameff C, Tully J, Chan TC, Castillo EM, Savage S, Maysent P, et al. Ransomware Attack Associated With Disruptions at Adjacent Emergency Departments in the US. JAMA Netw Open. 2023;6(5):e2312270. doi:10.1001/jamanetworkopen.2023.12270

185. Pham TT, Loo TM, Malhotra A, Longhurst CA, Hylton D, Dameff C, et al. Ransomware Cyberattack Associated With Cardiac Arrest Incidence and Outcomes at Untargeted, Adjacent Hospitals. Crit Care Explor. 2024;6(4):e1079. doi:10.1097/CCE.0000000000001079
186. Directoratul Național de Securitate Cibernetică. UPDATE: Un atac cibernetic de tip ransomware a afectat spitale din România [Internet]. București: DNSC; 2024 Feb 15 [updated 2024 Feb 16; cited 2026 Sep 27]. Available from: https://www.dnsc.ro/citeste/atac-cibernetic-ransomware-spitale-Romania. Archived at: http://web.archive.org/web/20240419005822/https://dnsc.ro/citeste/atac-cibernetic-ransomware-spitale-Romania
187. Ordonanța de urgență a Guvernului nr. 155/2024 privind instituirea unui cadru pentru securitatea cibernetică a rețelelor și sistemelor informatice din spațiul cibernetic național civil. Monitorul Oficial al României, Partea I. 2024 Dec 31;(1332).
188. Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive). Official Journal of the European Union. 2022 Dec 27;L 333:80.
189. ISO 27799:2025. Health informatics - Information security controls in health based on ISO/IEC 27002. 3rd ed. Geneva: International Organization for Standardization; 2025.
190. ISO/DIS 27799:2025. Health informatics - Information security management in health using ISO/IEC 27002. Draft International Standard (voting 2025 Jan 20 - 2025 Apr 14). Geneva: International Organization for Standardization; 2025.
191. Chen IY, Pierson E, Rose S, Joshi S, Ferryman K, Ghassemi M. Ethical Machine Learning in Healthcare. Annu Rev Biomed Data Sci. 2021;4:123-44. doi:10.1146/annurev-biodatasci-092820-114757
192. Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial Bias in Pulse Oximetry Measurement. N Engl J Med. 2020;383(25):2477-8. doi:10.1056/NEJMc2029240
193. Parr NJ, Beech EH, Young S, Valley TS. Racial and Ethnic Disparities in Occult Hypoxemia Prevalence and Clinical Outcomes Among Hospitalized Patients: A Systematic Review and Meta-analysis. J Gen Intern Med. 2024;39(13):2543-53. doi:10.1007/s11606-024-08852-1
194. Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. Nat Med. 2021;27(12):2176-82. doi:10.1038/s41591-021-01595-0
195. Inker LA, Eneanya ND, Coresh J, Tighiouart H, Wang D, Sang Y, et al. New Creatinine- and Cystatin C-Based Equations to Estimate GFR without Race. N Engl J Med. 2021;385(19):1737-49. doi:10.1056/NEJMoa2102953
196. Alderman JE, Palmer J, Laws E, McCradden MD, Ordish J, Ghassemi M, et al. Tackling algorithmic bias and promoting transparency in health datasets: the STANDING Together consensus recommendations. Lancet Digit Health. 2025;7(1):e64-88. doi:10.1016/S2589-7500(24)00224-3
197. Regulation (EU) 2026/1744 of the European Parliament and of the Council of 8 July 2026 amending Regulations (EU) 2024/1689, (EU) 2018/1139 and (EU) 2023/1230 as regards the simplification of the implementation of harmonised rules on artificial intelligence (Digital Omnibus on AI). Official Journal of the European Union. 2026 Jul 24;L 2026/1744.
198. Legea nr. 199/2023 a învățământului superior. Monitorul Oficial al României, Partea I. 2023 Jul 5;(614).
199. Fadlalmola HA, Elhusein AM, Swamy DSV, Hussein MK, Mamanao DM, Mohamedsalih WE. Plagiarism among nursing students: A systematic review and meta-analysis. Int Nurs Rev. 2022;69(4):492-502. doi:10.1111/inr.12755
200. Habibzadeh F. Plagiarism: A Bird's Eye View. J Korean Med Sci. 2023;38(45):e373. doi:10.3346/jkms.2023.38.e373
201. National Library of Medicine (US). Samples of Formatted References for Authors of Journal Articles [Internet]. Bethesda (MD): National Library of Medicine; [last reviewed 2026 Jun 2; cited 2026 Sep 27]. Available from: https://www.nlm.nih.gov/bsd/uniform_requirements.html
202. International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals [Internet]. [place unknown]: ICMJE; [updated 2026 Jan; cited 2026 Sep 27]. Available from: https://www.icmje.org/recommendations/
203. Committee on Publication Ethics. Authorship and AI tools. COPE position statement [Internet]. COPE; 2023 Feb 13 [cited 2026 Sep 27]. Available from: https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools
204. Ganjavi C, Eppler MB, Pekcan A, Biedermann B, Abreu A, Collins GS, et al. Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis. BMJ. 2024;384:e077192. doi:10.1136/bmj-2023-077192

205. Liang W, Yuksekgonul M, Mao Y, Wu E, Zou J. GPT detectors are biased against non-native English writers. Patterns (N Y). 2023;4(7):100779. doi:10.1016/j.patter.2023.100779
206. Herasevich V, Pickering BW, Clemmer TP, Mark RG. Patient Monitoring Systems. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 693-732. doi:10.1007/978-3-030-58721-5_21
207. Barmpagiannos K, Malik Z, Whyte S, Stavrakis S. Quantifying high-pass filter-induced ST-segment distortion. J Electrocardiol. 2026;99:154442. doi:10.1016/j.jelectrocard.2026.154442
208. Peltola ME, Leitinger M, Halford JJ, Vinayan KP, Kobayashi K, Pressler RM, et al. Routine and sleep EEG: Minimum recording standards of the International Federation of Clinical Neurophysiology and the International League Against Epilepsy. Epilepsia. 2023;64(3):602-18. doi:10.1111/epi.17448
209. Nayak CS, Anilkumar AC. Normal EEG Waveforms. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 [updated 2025 Aug 3; cited 2026 Sep 27]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK539805/
210. Sul JH, Piyathilaka L, Moratuwage D, Dunu Arachchige S, Jayawardena A, Kahandawa G, et al. Electromyography Signal Acquisition, Filtering, and Data Analysis for Exoskeleton Development. Sensors (Basel). 2025;25(13):4004. doi:10.3390/s25134004
211. Park J, Seok HS, Kim SS, Shin H. Photoplethysmogram Analysis and Applications: An Integrative Review. Front Physiol. 2021;12:808451. doi:10.3389/fphys.2021.808451
212. Quigley KS, Gianaros PJ, Norman GJ, Jennings JR, Berntson GG, de Geus EJC. Publication guidelines for human heart rate and heart rate variability studies in psychophysiology-Part 1: Physiological underpinnings and foundations of measurement. Psychophysiology. 2024;61(9):e14604. doi:10.1111/psyp.14604
213. PhysioNet. MIT-BIH Arrhythmia Database, version 1.0.0 [Internet]. Cambridge (MA): MIT Laboratory for Computational Physiology; 2026 [cited 2026 Sep 27]. Available from: https://physionet.org/content/mitdb/1.0.0/
214. Rahman S, Karmakar C, Natgunanathan I, Yearwood J, Palaniswami M. Robustness of electrocardiogram signal quality indices. J R Soc Interface. 2022;19(189):20220012. doi:10.1098/rsif.2022.0012
215. Babiloni C, Barry RJ, Başar E, Blinowska KJ, Cichocki A, Drinkenburg WHIM, et al. International Federation of Clinical Neurophysiology (IFCN) - EEG research workgroup: Recommendations on frequency and topographic analysis of resting state EEG rhythms. Part 1: Applications in clinical research studies. Clin Neurophysiol. 2020;131(1):285-307. doi:10.1016/j.clinph.2019.06.234
216. EDFbrowser, version 2.15. Free, open-source viewer for EDF/EDF+/BDF files [Internet]. Teuniz.net; 2026 [cited 2026 Sep 27]. Available from: https://www.teuniz.net/edfbrowser/
217. Weinstein A, Rodiño J, Otero M. Effects of electrocardiogram qrs detection algorithms in heart rate variability metrics. Sci Rep. 2026;16(1):18589. doi:10.1038/s41598-026-49215-6
218. Littmann L. Electrocardiographic artifact. J Electrocardiol. 2021;64:23-9. doi:10.1016/j.jelectrocard.2020.11.006
219. Amin U, Nascimento FA, Karakis I, Schomer D, Benbadis SR. Normal variants and artifacts: Importance in EEG interpretation. Epileptic Disord. 2023;25(5):591-648. doi:10.1002/epd2.20040
220. van der Bijl K, Elgendi M, Menon C. Automatic ECG Quality Assessment Techniques: A Systematic Review. Diagnostics (Basel). 2022;12(11):2578. doi:10.3390/diagnostics12112578
221. European Data Format. EDF and EDF+ [Internet]. 2026 [cited 2026 Sep 27]. Available from: https://www.edfplus.info/
222. WFDB: Waveform Database. Open standards and software for physiologic signal and annotation data [Internet]. 2026 [cited 2026 Sep 27]. Available from: https://wfdb.io/
223. National Electrical Manufacturers Association. DICOM PS3.3 2026d - Information Object Definitions. Section A.34: Waveform IODs [Internet]. Rosslyn (VA): NEMA; 2026 [cited 2026 Sep 27]. Available from: https://dicom.nema.org/medical/dicom/current/output/chtml/part03/sect_A.34.html
224. Health Level Seven International. HL7 Version 3 Implementation Guide: Annotated ECG (aECG) R1, Release 2 - US Realm. Product brief [Internet]. Ann Arbor (MI): HL7 International; c2007-2026 [cited 2026 Sep 27]. Available from: https://www.hl7.org/implement/standards/product_brief.cfm?product_id=102
225. ISO/IEEE 11073-10103:2025. Health informatics - Device interoperability - Part 10103: Nomenclature - Implantable device, cardiac. Geneva: International Organization for Standardization; 2025.
226. PhysioNet. The research resource for complex physiologic signals [Internet]. Cambridge (MA): MIT Laboratory for Computational Physiology; 2026 [cited 2026 Sep 27]. Available from: https://physionet.org/

227. Pollard T, Moody BE, Lehman LwH, Gow BJ, Fernandes C, Xie C, et al. PhysioNet as a global platform for biomedical research. Nat Health. 2026;1(8):792-5. doi:10.1038/s44360-026-00096-z
228. PhysioNet. LightWAVE 0.71: waveform and annotation viewer [Internet]. Cambridge (MA): MIT Laboratory for Computational Physiology; 2026 [cited 2026 Sep 27]. Available from: https://physionet.org/lightwave/
229. Ose B, Sattar Z, Gupta A, Toquica C, Harvey C, Noheria A. Artificial Intelligence Interpretation of the Electrocardiogram: A State-of-the-Art Review. Curr Cardiol Rep. 2024;26(6):561-80. doi:10.1007/s11886-024-02062-1
230. Yao X, Rushlow DR, Inselman JW, McCoy RG, Thacher TD, Behnken EM, et al. Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. Nat Med. 2021;27(5):815-9. doi:10.1038/s41591-021-01335-4
231. Tsai DJ, Lin C, Liu WT, Lee CC, Chang CH, Lin WY, et al. Artificial intelligence-assisted diagnosis and prognostication in low ejection fraction using electrocardiograms in inpatient department: a pragmatic randomized controlled trial. BMC Med. 2025;23(1):342. doi:10.1186/s12916-025-04190-z
232. Adedinsewo DA, Morales-Lara AC, Afolabi BB, Kushimo OA, Mbakwem AC, Ibiyemi KF, et al. Artificial intelligence guided screening for cardiomyopathies in an obstetric population: a pragmatic randomized clinical trial. Nat Med. 2024;30(10):2897-906. doi:10.1038/s41591-024-03243-9
233. Christopoulos G, Graff-Radford J, Lopez CL, Yao X, Attia ZI, Rabinstein AA, et al. Artificial Intelligence-Electrocardiography to Predict Incident Atrial Fibrillation: A Population-Based Study. Circ Arrhythm Electrophysiol. 2020;13(12):e009355. doi:10.1161/CIRCEP.120.009355
234. Charlton PH, Allen J, Bailón R, Baker S, Behar JA, Chen F, et al. The 2023 wearable photoplethysmography roadmap. Physiol Meas. 2023;44(11):111001. doi:10.1088/1361-6579/acead2
235. Lubitz SA, Faranesh AZ, Selvaggi C, Atlas SJ, McManus DD, Singer DE, et al. Detection of Atrial Fibrillation in a Large Population Using Wearable Devices: The Fitbit Heart Study. Circulation. 2022;146(19):1415-24. doi:10.1161/CIRCULATIONAHA.122.060291
236. Lubitz SA, McConnell MV, Selvaggi C, Krishnamoorthy A, Atlas SJ, McManus DD, et al. Wearable Irregular Heart Rhythm Detection Recurrences and Electrocardiographic Atrial Fibrillation Confirmation: The Fitbit Heart Study. Circ Arrhythm Electrophysiol. 2025;18(7):e013565. doi:10.1161/CIRCEP.124.013565
237. van Steijn NJ, Blommestijn IS, Blok S, Pepplinkhuizen S, Somsen GA, Knops RE, et al. Enhanced Detection and Prompt Diagnosis of Atrial Fibrillation Using Apple Watch: A Randomized Controlled Trial. J Am Coll Cardiol. 2026;87(14):1714-28. doi:10.1016/j.jacc.2025.11.032
238. Barrera N, Solorzano M, Jimenez Y, Kushnir Y, Gallegos-Koyner F, Dagostin de Carvalho G. Accuracy of Smartwatches in the Detection of Atrial Fibrillation: A Systematic Review and Diagnostic Meta-Analysis. JACC Adv. 2025;4(11 Pt 1):102133. doi:10.1016/j.jacadv.2025.102133
239. Brandes A, Stavrakis S, Freedman B, Antoniou S, Boriani G, Camm AJ, et al. Consumer-Led Screening for Atrial Fibrillation: Frontier Review of the AF-SCREEN International Collaboration. Circulation. 2022;146(19):1461-74. doi:10.1161/CIRCULATIONAHA.121.058911
240. Haase CB, Jensen AE, Modin FA, Siersma V, Brodersen JB. Overdiagnosis in atrial fibrillation screening with wearables. Scand J Prim Health Care. 2026;44(1):2656694. doi:10.1080/02813432.2026.2656694
241. Noseworthy PA, Attia ZI, Brewer LC, Hayes SN, Yao X, Kapa S, et al. Assessing and Mitigating Bias in Medical Artificial Intelligence: The Effects of Race and Ethnicity on a Deep Learning Model for ECG Analysis. Circ Arrhythm Electrophysiol. 2020;13(3):e007988. doi:10.1161/CIRCEP.119.007988
242. Singh S, Bennett MR, Chen C, Shin S, Ghanbari H, Nelson BW. Impact of Skin Pigmentation on Pulse Oximetry Blood Oxygenation and Wearable Pulse Rate Accuracy: Systematic Review and Meta-Analysis. J Med Internet Res. 2024;26:e62769. doi:10.2196/62769
243. Muntean C, Gaborean V, Faur AM, Faur IF, Prodan-Bărbulescu C, Feier CVI. Continuous Wearable-Sensor Monitoring After Colorectal Surgery: A Systematic Review of Clinical Outcomes and Predictive Analytics. Diagnostics (Basel). 2025;15(17):2194. doi:10.3390/diagnostics15172194
244. Ministerul Sănătății. Caiet de sarcini pentru servicii de dezvoltare și implementare sistem informatic integrat pentru evidența clinică a secțiilor A.T.I. (S.I.E.C.-A.T.I.) [Internet]. București: Ministerul Sănătății; 2021 [cited 2026 Sep 27]. Available from: https://ms.ro/media/documents/Caiet-de-Sarcini-Sist-integrat-SIEC-ATI.pdf
245. National Electrical Manufacturers Association. Digital Imaging and Communications in Medicine (DICOM) Standard, PS3.1-PS3.22, 2026d edition [Internet]. Rosslyn (VA): NEMA; 2026 [cited 2026 Sep 27]. Available from: https://www.dicomstandard.org/current
246. Dhamija E, Gulati M, Deo SVS, Gogia A, Hari S. Digital Breast Tomosynthesis: an Overview. Indian J Surg Oncol. 2021;12(2):315-29. doi:10.1007/s13193-021-01310-y

247. Cornish TC, Dayal S, Ferber D, Chiweshe J, Monroe R. The evolving role of DICOM in digital pathology. J Pathol Inform. 2026;21:100658. doi:10.1016/j.jpi.2026.100658
248. Evans AJ, Brown RW, Bui MM, Chlipala EA, Lacchetti C, Milner DA, et al. Validating Whole Slide Imaging Systems for Diagnostic Purposes in Pathology. Arch Pathol Lab Med. 2022;146(4):440-50. doi:10.5858/arpa.2020-0723-CP
249. Jo YY, Choi YS, Park HW, Lee JH, Jung H, Kim HE, et al. Impact of image compression on deep learning-based mammogram classification. Sci Rep. 2021;11(1):7924. doi:10.1038/s41598-021-86726-w
250. Trieu PDY, Barron M, Lewis SJ. Use of Full-quality DICOM Images Compared to Minimally Compressed Mammograms in JPEG Format for Radiology Training: A Study From Radiologist and Radiographer Perspectives. Acad Radiol. 2023;30(8):1748-55. doi:10.1016/j.acra.2022.11.012
251. Fischer M, Neher P, Schüffler P, Ziegler S, Xiao S, Peretzke R, et al. Unlocking the potential of digital pathology: Novel baselines for compression. J Pathol Inform. 2025;17:100421. doi:10.1016/j.jpi.2025.100421
252. ISO/DIS 12052:2025. Health informatics - Digital imaging and communication in medicine (DICOM) including workflow and data management. Draft International Standard. Geneva: International Organization for Standardization; 2025.
253. ISO 12052:2026. Health informatics - Digital imaging and communication in medicine (DICOM) including workflow and data management. 3rd ed. Geneva: International Organization for Standardization; 2026.
254. National Electrical Manufacturers Association. DICOMweb [Internet]. Rosslyn (VA): NEMA; 2026 [cited 2026 Sep 27]. Available from: https://www.dicomstandard.org/using/dicomweb
255. IHE International. Scheduled Workflow (SWF) [Internet]. IHE Wiki; [cited 2026 Sep 27]. Available from: https://wiki.ihe.net/index.php/Scheduled_Workflow
256. Tang ST, Tjia V, Noga T, Febri J, Lien CY, Chu WC, et al. Creating a Medical Imaging Workflow Based on FHIR, DICOMweb, and SVG. J Digit Imaging. 2023;36(3):794-803. doi:10.1007/s10278-021-00522-6
257. Clunie DA, Napora B, Groth J, Zerbe N, Yousif M, Shelar G, et al. Report on the 2025 DICOM WSI Connectathon. J Pathol Inform. 2026;21:100657. doi:10.1016/j.jpi.2026.100657
258. Clunie D, Prior F, Rutherford M, Moore S, Parker W, Kondylakis H, et al. Summary of the National Cancer Institute 2023 Virtual Workshop on Medical Image De-identification-Part 1: Report of the MIDI Task Group - Best Practices and Recommendations, Tools for Conventional Approaches to De-identification, International Approaches to De-identification, and Industry Panel on Image De-identification. J Imaging Inform Med. 2025;38(1):1-15. doi:10.1007/s10278-024-01182-y
259. Eichelberg M, Kleber K, Kämmerer M. Cybersecurity in PACS and Medical Imaging: an Overview. J Digit Imaging. 2020;33(6):1527-42. doi:10.1007/s10278-020-00393-3
260. Ghaderi H, April A. Assessing Available Open-Source PACS Options. J Digit Imaging. 2023;36(6):2323-8. doi:10.1007/s10278-021-00435-4
261. IHE International. IHE Radiology (RAD) Technical Framework, Revision 23.0 [Internet]. IHE International; 2025 Aug 8 [cited 2026 Sep 27]. Available from: https://profiles.ihe.net/RAD/index.html
262. Kawa J, Pyciński B, Smoliński M, Bożek P, Kwasecki M, Pietrzyk B, et al. Design and Implementation of a Cloud PACS Architecture. Sensors (Basel). 2022;22(21):8569. doi:10.3390/s22218569
263. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods. 2021;18(2):203-11. doi:10.1038/s41592-020-01008-z
264. Wasserthal J, Breit HC, Meyer MT, Pradella M, Hinck D, Sauter AW, et al. TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. Radiol Artif Intell. 2023;5(5):e230024. doi:10.1148/ryai.230024
265. van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging-"how-to" guide and critical reflection. Insights Imaging. 2020;11(1):91. doi:10.1186/s13244-020-00887-2
266. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. Radiology. 2020;295(2):328-38. doi:10.1148/radiol.2020191145
267. U.S. Food and Drug Administration. Artificial Intelligence-Enabled Medical Devices [Internet]. Silver Spring (MD): FDA; 2026 Sep 4 [cited 2026 Sep 27]. Available from: https://www.fda.gov/medical-devices/artificial-intelligence-enabled-medical-devices/list-artificial-intelligence-enabled-medical-devices
268. Lee YS, Youn BY. Clinical Specialty Expansion of AI-Enabled and Machine Learning-Enabled Medical Devices Authorized by the US Food and Drug Administration From 1995 to 2025: Longitudinal Content Analysis. J Med Internet Res. 2026;28:e103040. doi:10.2196/103040

269. Lång K, Josefsson V, Larsson AM, Larsson S, Högberg C, Sartor H, et al. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. Lancet Oncol. 2023;24(8):936-44. doi:10.1016/S1470-2045(23)00298-X
270. Hernström V, Josefsson V, Sartor H, Schmidt D, Larsson AM, Hofvind S, et al. Screening performance and characteristics of breast cancer detected in the Mammography Screening with Artificial Intelligence trial (MASAI): a randomised, controlled, parallel-group, non-inferiority, single-blinded, screening accuracy study. Lancet Digit Health. 2025;7(3):e175-83. doi:10.1016/S2589-7500(24)00267-X
271. Gommers J, Hernström V, Josefsson V, Sartor H, Schmidt D, Hjelmgren A, et al. Interval cancer, sensitivity, and specificity comparing AI-supported mammography screening with standard double reading without AI in the MASAI study: a randomised, controlled, non-inferiority, single-blinded, population-based, screening-accuracy trial. Lancet. 2026;407(10527):505-14. doi:10.1016/S0140-6736(25)02464-X
272. Eisemann N, Bunk S, Mukama T, Baltus H, Elsner SA, Gomille T, et al. Nationwide real-world implementation of AI for cancer detection in population-based mammography screening. Nat Med. 2025;31(3):917-24. doi:10.1038/s41591-024-03408-6
273. Geppert J, Asgharzadeh A, Brown A, Stinton C, Helm EJ, Jayakody S, et al. Software using artificial intelligence for nodule and cancer detection in CT lung cancer screening: systematic review of test accuracy studies. Thorax. 2024;79(11):1040-9. doi:10.1136/thorax-2024-221662
274. Geppert J, Auguste P, Asgharzadeh A, Ghiasvand H, Patel M, Brown A, et al. Software with artificial intelligence-derived algorithms for detecting and analysing lung nodules in CT scans: systematic review and economic evaluation. Health Technol Assess. 2025;29(14):1-234. doi:10.3310/JYTW8921
275. Yu AC, Mohajer B, Eng J. External Validation of Deep Learning Algorithms for Radiologic Diagnosis: A Systematic Review. Radiol Artif Intell. 2022;4(3):e210064. doi:10.1148/ryai.210064
276. Kelly BS, Judge C, Bollard SM, Clifford SM, Healy GM, Aziz A, et al. Radiology artificial intelligence: a systematic review and evaluation of methods (RAISE). Eur Radiol. 2022;32(11):7998-8007. doi:10.1007/s00330-022-08784-6
277. McGenity C, Clarke EL, Jennings C, Matthews G, Cartlidge C, Freduah-Agyemang H, et al. Artificial intelligence in digital pathology: a systematic review and meta-analysis of diagnostic test accuracy. NPJ Digit Med. 2024;7(1):114. doi:10.1038/s41746-024-01106-8
278. Komura D, Ochi M, Ishikawa S. Machine learning methods for histopathological image analysis: Updates in 2024. Comput Struct Biotechnol J. 2025;27:383-400. doi:10.1016/j.csbj.2024.12.033
279. Varoquaux G, Cheplygina V. Machine learning for medical imaging: methodological failures and recommendations for the future. NPJ Digit Med. 2022;5(1):48. doi:10.1038/s41746-022-00592-y
280. Tejani AS, Ng YS, Xi Y, Rayan JC. Understanding and Mitigating Bias in Imaging Artificial Intelligence. Radiographics. 2024;44(5):e230067. doi:10.1148/rg.230067
281. Dratsch T, Chen X, Rezazade Mehrizi M, Kloeckner R, Mähringer-Kunz A, Püsken M, et al. Automation Bias in Mammography: The Impact of Artificial Intelligence BI-RADS Suggestions on Reader Performance. Radiology. 2023;307(4):e222176. doi:10.1148/radiol.222176
282. Bernstein MH, Atalay MK, Dibble EH, Maxwell AWP, Karam AR, Agarwal S, et al. Can incorrect artificial intelligence (AI) results impact radiologists, and if so, what can we do about it? A multi-reader pilot study of lung cancer detection with chest radiography. Eur Radiol. 2023;33(11):8263-9. doi:10.1007/s00330-023-09747-1
283. Yu F, Moehring A, Banerjee O, Salz T, Agarwal N, Rajpurkar P. Heterogeneity and predictors of the effects of AI assistance on radiologists. Nat Med. 2024;30(3):837-49. doi:10.1038/s41591-024-02850-w
284. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. Official Journal of the European Union. 2017 May 5;L 117:1.
285. Melton GB, McDonald CJ, Tang PC, Hripcsak G. Electronic Health Records. In: Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 5th ed. Cham: Springer; 2021. p. 467-509. doi:10.1007/978-3-030-58721-5_14
286. Brands MR, Gouw SC, Beestrum M, Cronin RM, Fijnvandraat K, Badawy SM. Patient-Centered Digital Health Records and Their Effects on Health Outcomes: Systematic Review. J Med Internet Res. 2022;24(12):e43086. doi:10.2196/43086
287. Wright A, Schreiber R, Bates DW, Aaron S, Ai A, Cholan RA, et al. A multi-site randomized trial of a clinical decision support intervention to improve problem list completeness. J Am Med Inform Assoc. 2023;30(5):899-906. doi:10.1093/jamia/ocad020

288. Tsang JY, Planner C, Low CN, Stewart S, Morris C, Wong B, et al. Digital patient safety interventions in primary care: a systematic review and meta-analysis. BMC Med. 2026;24(1):407. doi:10.1186/s12916-026-05048-8
289. Pinevich Y, Clark KJ, Harrison AM, Pickering BW, Herasevich V. Interaction Time with Electronic Health Records: A Systematic Review. Appl Clin Inform. 2021;12(4):788-99. doi:10.1055/s-0041-1733909
290. Overhage JM, McCallie D. Physician Time Spent Using the Electronic Health Record During Outpatient Encounters: A Descriptive Study. Ann Intern Med. 2020;172(3):169-74. doi:10.7326/M18-3684
291. Holmgren AJ, Downing NL, Bates DW, Shanafelt TD, Milstein A, Sharp CD, et al. Assessment of Electronic Health Record Use Between US and Non-US Health Systems. JAMA Intern Med. 2021;181(2):251-9. doi:10.1001/jamainternmed.2020.7071
292. Arndt BG, Micek MA, Rule A, Shafer CM, Baltus JJ, Sinsky CA. More Tethered to the EHR: EHR Workload Trends Among Academic Primary Care Physicians, 2019-2023. Ann Fam Med. 2024;22(1):12-8. doi:10.1370/afm.3047
293. Moy AJ, Schwartz JM, Elias J, Imran S, Lucas E, Cato KD, et al. Time-motion examination of electronic health record utilization and clinician workflows indicate frequent task switching and documentation burden. AMIA Annu Symp Proc. 2020;2020:886-95.
294. Wu Y, Wu M, Wang C, Lin J, Liu J, Liu S. Evaluating the Prevalence of Burnout Among Health Care Professionals Related to Electronic Health Record Use: Systematic Review and Meta-Analysis. JMIR Med Inform. 2024;12:e54811. doi:10.2196/54811
295. Holmgren AJ, Hendrix N, Maisel N, Everson J, Bazemore A, Rotenstein L, et al. Electronic Health Record Usability, Satisfaction, and Burnout for Family Physicians. JAMA Netw Open. 2024;7(8):e2426956. doi:10.1001/jamanetworkopen.2024.26956
296. Rotenstein LS, Holmgren AJ, Thombley R, Sriram A, Dbouk RH, Jost M, et al. Changes in Clinician Time Expenditure and Visit Quantity With Adoption of Artificial Intelligence-Powered Scribes: A Multisite Study. JAMA. 2026;335(16):1408-17. doi:10.1001/jama.2026.2253
297. Eurostat. Individuals - internet activities (isoc_ci_ac_i) [Internet]. Luxembourg: Eurostat; 2026 Apr 17 [cited 2026 Sep 27]. doi:10.2908/ISOC_CI_AC_I. Available from: https://ec.europa.eu/eurostat/databrowser/view/isoc_ci_ac_i/default/table?lang=en
298. El-Yafouri R, Klieb L. A scoping review of electronic health records interoperability levels, expectations, approaches, and problems. Health Informatics J. 2025;31(4):14604582251385986. doi:10.1177/14604582251385986
299. Luu HS, Campbell WS, Cholan RA, Edgerton ME, Englund A, Keller A, et al. Analysis of laboratory data transmission between two healthcare institutions using a widely used point-to-point health information exchange platform: a case report. JAMIA Open. 2024;7(2):ooae032. doi:10.1093/jamiaopen/ooae032
300. HL7 International. HL7 Version 2 Product Suite [Internet]. Ann Arbor (MI): HL7 International; 2026 [cited 2026 Sep 27]. Available from: https://www.hl7.org/implement/standards/product_brief.cfm?product_id=185
301. HL7 International. HL7 Clinical Document Architecture (CDA) Release 2 [Internet]. Ann Arbor (MI): HL7 International; 2026 [cited 2026 Sep 27]. Available from: https://www.hl7.org/implement/standards/product_brief.cfm?product_id=7
302. openEHR International. What is openEHR [Internet]. openEHR International; 2026 [cited 2026 Sep 27]. Available from: https://openehr.org/about/what_is_openehr
303. Pedrera-Jiménez M, García-Barrio N, Frid S, Moner D, Boscá-Tomás D, Lozano-Rubí R, et al. Can OpenEHR, ISO 13606, and HL7 FHIR Work Together? An Agnostic Approach for the Selection and Application of Electronic Health Record Standards to the Next-Generation Health Data Spaces. J Med Internet Res. 2023;25:e48702. doi:10.2196/48702
304. Marfoglia A, Arcobelli VA, Moscato S, La Mattina AA, Mellone S, Carbonaro A. Challenges of health data standard adoption and usage: a systematic review. J Biomed Inform. 2026;178:105022. doi:10.1016/j.jbi.2026.105022
305. HL7 International. FHIR Release 5 (v5.0.0) [Internet]. Ann Arbor (MI): HL7 International; 2023 [cited 2026 Sep 27]. Available from: https://hl7.org/fhir/R5/
306. HL7 International. SMART App Launch Implementation Guide, release 2.2.0 (STU 2.2) [Internet]. Ann Arbor (MI): HL7 International; 2024 [cited 2026 Sep 27]. Available from: https://hl7.org/fhir/smart-app-launch/
307. HL7 International. FHIR publication (version) history [Internet]. Ann Arbor (MI): HL7 International; 2026 [cited 2026 Sep 27]. Available from: https://hl7.org/fhir/directory.html

308. Wesley DB, Blumenthal J, Shah S, Littlejohn RA, Pruitt Z, Dixit R, et al. A novel application of SMART on FHIR architecture for interoperable and scalable integration of patient-reported outcome data with electronic health records. J Am Med Inform Assoc. 2021;28(10):2220-5. doi:10.1093/jamia/ocab110
309. HL7 International. CDS Hooks Specification, STU 2 Release (v2.0.1) [Internet]. Ann Arbor (MI): HL7 International; 2025 [cited 2026 Sep 27]. Available from: https://cds-hooks.hl7.org/
310. Morgan KL, Kukhareva PV, Warner PB, Wilkof J, Snyder M, Horton D, et al. Using CDS Hooks to increase SMART on FHIR app utilization: a cluster-randomized trial. J Am Med Inform Assoc. 2022;29(9):1461-70. doi:10.1093/jamia/ocac085
311. ISO 27269:2025. Health informatics - International patient summary. 2nd ed. Geneva: International Organization for Standardization; 2025.
312. ISO 17523:2025. Health informatics - Requirements for electronic prescriptions. 2nd ed. Geneva: International Organization for Standardization; 2025.
313. HL7 International. International Patient Summary Implementation Guide, release 2.0.1 (STU 2) [Internet]. Ann Arbor (MI): HL7 International; 2026 [cited 2026 Sep 27]. Available from: https://hl7.org/fhir/uv/ips/
314. D'Amore JD, Vallapu ER, Samal L, Karras BT, Lober WB. From Pandemic Response to Kill the Clipboard: Patient-Controlled Sharing of Health Data Using International Patient Summary (IPS) and QR codes. Appl Clin Inform. 2026;17(3):559-67. doi:10.1055/a-2902-8515
315. Laleci Erturkmen GB, Sinaci AA, Namli T, Boonstra MJ, Lekadir K, Gkontra P, et al. An assessment of the European Patient Summary for clinical research: a case study in cardiology. Front Med (Lausanne). 2024;11:1481551. doi:10.3389/fmed.2024.1481551
316. European Commission. Electronic cross-border health services (MyHealth@EU) [Internet]. Brussels: European Commission; 2026 [cited 2026 Sep 27]. Available from: https://health.ec.europa.eu/ehealth-digital-health-and-care/electronic-cross-border-health-services_en
317. Bruthans J, Jiráková K. The Current State and Usage of European Electronic Cross-border Health Services (eHDSI). J Med Syst. 2023;47(1):21. doi:10.1007/s10916-023-01920-9
318. Jõgi R, Timonen J, Saastamoinen L, Laius O, Volmer D. Implementation of European Cross-border Electronic Prescription and Electronic Dispensing Service: Cross-sectional Survey. J Med Internet Res. 2023;25:e42453. doi:10.2196/42453
319. European Commission. European Health Data Space Regulation (EHDS) [Internet]. Brussels: European Commission; 2026 [cited 2026 Sep 27]. Available from: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en
320. Donia J, Marelli L. Anticipating ethical and social dimensions of the European Health Data Space: A rapid systematic review. Health Policy. 2025;162:105443. doi:10.1016/j.healthpol.2025.105443
321. Legea nr. 45/2019 pentru modificarea şi completarea Legii nr. 95/2006 privind reforma în domeniul sănătăţii. Monitorul Oficial al României, Partea I. 2019 Mar 11;(192).
322. Casa Națională de Asigurări de Sănătate. Portal CNAS-PIAS: Platforma Integrată a Asigurărilor de Sănătate [Internet]. Bucureşti: CNAS; 2026 [cited 2026 Sep 27]. Available from: https://portal.cnas.ro/cnas/
323. Casa Națională de Asigurări de Sănătate. Comunicat de presă: Finalizare proiect "Redimensionarea, standardizarea şi optimizarea Platformei Informatice din Asigurările de Sănătate - PIAS" [Internet]. Bucureşti: CNAS; 2026 Aug 31 [cited 2026 Sep 27]. Available from: https://cnas.ro/wp-content/uploads/2026/09/Comunicat-PNRR-01092026b-final.pdf
324. Casa Națională de Asigurări de Sănătate. Comunicat: CNAS a lansat portalul eSănătateaMea [Internet]. Bucureşti: CNAS; 2026 Sep 2 [cited 2026 Sep 27]. Available from: https://cnas.ro/2026/09/02/comunicat-cnas-a-lansat-portalul-esanatateamea/
325. Casa Națională de Asigurări de Sănătate. Comunicat: Precizări privind utilizarea CEAS/CEI în sistemul asigurărilor sociale de sănătate [Internet]. Bucureşti: CNAS; 2026 Aug 31 [cited 2026 Sep 27]. Available from: https://cnas.ro/2026/08/31/comunicat-precizari-privind-utilizarea-ceas-cei-in-sistemul-asigurarilor-sociale-de-sanatate/
326. Ministerul Sănătății. Componenta 7 - Transformare digitală. Investiția specifică I3.3 - Investiții în sistemele informatice şi în infrastructura digitală a unităților sanitare publice [Internet]. Bucureşti: Ministerul Sănătății; 2023 [cited 2026 Sep 27]. Available from: https://ms.ro/ro/de-interes/apeluri-pnrr-c12-sanatate/componenta-7-transformare-digital%C4%83-investitia-specific%C4%83-i33-investi%C8%9Bii-%C3%AEn-sistemele-informatice-%C8%99i-%C3%AEn-infrastructura-digital%C4%83-a-unit%C4%83%C8%9Bilor-sanitare-publice/
327. Spitalul Clinic Județean de Urgență "Pius Brînzeu" Timişoara. Sănătatea, siguranța și confortul pacientului prin servicii integrate digitale la SCJU Timişoara - PNRR, Componenta C7 - Transformare digitală, PNRR/C7/MS-733/I3.3 [Internet]. Timişoara: SCJUPBT; 2024 Aug 17 [cited 2026 Sep 27]. Available

from: https://www.hosptm.ro/sanatatea-siguranta-si-confortul-pacientului-prin-servicii-integrate-digitale-la-scju-timisoara-pnrr-componenta-c7-transformare-digitala-pnrr-c7-ms-733-i3-3/

328. Adams KT, Pruitt Z, Kazi S, Hettinger AZ, Howe JL, Fong A, et al. Identifying Health Information Technology Usability Issues Contributing to Medication Errors Across Medication Process Stages. J Patient Saf. 2021;17(8):e988-94. doi:10.1097/PTS.0000000000000868
329. Schwappach D, Hautz W, Krummrey G, Pfeiffer Y, Ratwani RM. EMR usability and patient safety: a national survey of physicians. NPJ Digit Med. 2025;8(1):282. doi:10.1038/s41746-025-01657-4
330. Edrees H, Amato MG, Wong A, Seger DL, Bates DW. High-priority drug-drug interaction clinical decision support overrides in a newly implemented commercial computerized provider order-entry system: Override appropriateness and adverse drug events. J Am Med Inform Assoc. 2020;27(6):893-900. doi:10.1093/jamia/ocaa034
331. Ray CE, Wilson GM, Hughes AM, Cunningham Goedken C, Liu EP, Fitzpatrick MA, et al. Alert fatigue measurement in clinical decision support: a systematic review. J Am Med Inform Assoc. 2026;33(8):1523-31. doi:10.1093/jamia/ocag064
332. Steinkamp J, Kantrowitz JJ, Airan-Javia S. Prevalence and Sources of Duplicate Information in the Electronic Medical Record. JAMA Netw Open. 2022;5(9):e2233348. doi:10.1001/jamanetworkopen.2022.33348
333. Sittig DF, Flanagan T, Sengstack P, Cholankeril RT, Ehsan S, Heidemann A, et al. Revisions to the Safety Assurance Factors for Electronic Health Record Resilience (SAFER) Guides to update national recommendations for safe use of electronic health records. J Am Med Inform Assoc. 2025;32(4):755-60. doi:10.1093/jamia/ocaf018
334. World Health Organization. Consolidated telemedicine implementation guide [Internet]. Geneva: WHO; 2022 [cited 2026 Sep 27]. Available from: https://www.who.int/publications/i/item/9789240059184
335. World Health Organization. Classification of digital interventions, services and applications in health: a shared language to describe the uses of digital technology for health. 2nd ed [Internet]. Geneva: WHO; 2023 [cited 2026 Sep 27]. Available from: https://www.who.int/publications/i/item/9789240081949
336. O'Shea AMJ, Howren MB, Mulligan K, Haraldsson B, Shahnazi A, Kaboli PJ. Quantifying the Digital Divide: Associations of Broadband Internet with Tele-mental Health Access Before and During the COVID-19 Pandemic. J Gen Intern Med. 2023;38(Suppl 3):832-40. doi:10.1007/s11606-023-08120-8
337. Fan M, Wang Q, Liu J, Su L, Wang B, Xu H, et al. Real-world evaluation of the Stemoscope electronic tele-auscultation system. Biomed Eng Online. 2022;21(1):63. doi:10.1186/s12938-022-01032-4
338. Martyin K, Meznerics FA, Bokor LA, Szabó B, Hegyi P, Kiss N, et al. Diagnostic accuracy of teledermatology for skin diseases: a systematic review and meta-analysis. Front Med (Lausanne). 2026;13:1739592. doi:10.3389/fmed.2026.1739592
339. Tong SYK, Jackson TM, Lau AYS. Virtual physical examination in teleconsultation: A scoping review. Int J Med Inform. 2024;191:105561. doi:10.1016/j.ijmedinf.2024.105561
340. Jitsi. Security and privacy in Jitsi Meet [Internet]. 2026 [cited 2026 Sep 27]. Available from: https://jitsi.org/security/
341. Tukur M, Saad G, AlShagathrh FM, Househ M, Agus M. Telehealth interventions during COVID-19 pandemic: a scoping review of applications, challenges, privacy and security issues. BMJ Health Care Inform. 2023;30(1):e100676. doi:10.1136/bmjhci-2022-100676
342. Demaerschalk BM, Pines A, Butterfield R, Haglin JM, Haddad TC, Yiannias J, et al. Assessment of Clinician Diagnostic Concordance With Video Telemedicine in the Integrated Multispecialty Practice at Mayo Clinic During the Beginning of COVID-19 Pandemic From March to June 2020. JAMA Netw Open. 2022;5(9):e2229958. doi:10.1001/jamanetworkopen.2022.29958
343. Patel SY, Mehrotra A, Huskamp HA, Uscher-Pines L, Ganguli I, Barnett ML. Variation In Telemedicine Use And Outpatient Care During The COVID-19 Pandemic In The United States. Health Aff (Millwood). 2021;40(2):349-58. doi:10.1377/hlthaff.2020.01786
344. Kirk UB, Pellin MCG, Hoffmann K, Stummer FO, Van Poel E, Vestergaard CH, et al. Trends in the use of video consultation in general practice during COVID-19: impact of practice and country characteristics based on the international, cross-sectional PRICOV-19 study. BMC Prim Care. 2025;26(1):205. doi:10.1186/s12875-025-02892-0
345. McBain RK, Schuler MS, Breslau J, Kofner A, Wang L, Cantor JH. Telehealth Availability for Mental Health Care During and After the COVID-19 Public Health Emergency. JAMA Netw Open. 2024;7(7):e2420853. doi:10.1001/jamanetworkopen.2024.20853
346. Scholte NTB, Gürgöze MT, Aydin D, Theuns DAMJ, Manintveld OC, Ronner E, et al. Telemonitoring for heart failure: a meta-analysis. Eur Heart J. 2023;44(31):2911-26. doi:10.1093/eurheartj/ehad280

347. Kuan PX, Chan WK, Fern Ying DK, Rahman MAA, Peariasamy KM, Lai NM, et al. Efficacy of telemedicine for the management of cardiovascular disease: a systematic review and meta-analysis. Lancet Digit Health. 2022;4(9):e676-91. doi:10.1016/S2589-7500(22)00124-8
348. Kaihara T, Intan-Goey V, Scherrenberg M, Falter M, Kario K, Akashi Y, et al. Automatic transmission of home blood pressure data can be effective in managing hypertension: a systematic review and meta-analysis. Eur Heart J Digit Health. 2022;3(4):638-53. doi:10.1093/ehjdh/ztac049
349. Grover S, Mishra HP, Gupta R, Gupta LK. Effect of telemonitoring and home blood pressure monitoring on blood pressure reduction in hypertensive adults: a network meta-analysis. J Hypertens. 2025;43(7):1091-8. doi:10.1097/HJH.0000000000004008
350. Hangaard S, Laursen SH, Andersen JD, Kronborg T, Vestergaard P, Hejlesen O, et al. The Effectiveness of Telemedicine Solutions for the Management of Type 2 Diabetes: A Systematic Review, Meta-Analysis, and Meta-Regression. J Diabetes Sci Technol. 2023;17(3):794-825. doi:10.1177/19322968211064633
351. Zhang J, Ji X, Xie J, Lin K, Yao M, Chi C. Effectiveness of synchronous teleconsultation for patients with type 2 diabetes mellitus: a systematic review and meta-analysis. BMJ Open Diabetes Res Care. 2023;11(1):e003180. doi:10.1136/bmjdrc-2022-003180
352. Janjua S, Carter D, Threapleton CJ, Prigmore S, Disler RT. Telehealth interventions: remote monitoring and consultations for people with chronic obstructive pulmonary disease (COPD). Cochrane Database Syst Rev. 2021;7(7):CD013196. doi:10.1002/14651858.CD013196.pub2
353. Cox NS, Dal Corso S, Hansen H, McDonald CF, Hill CJ, Zanaboni P, et al. Telerehabilitation for chronic respiratory disease. Cochrane Database Syst Rev. 2021;1(1):CD013040. doi:10.1002/14651858.CD013040.pub2
354. Hagi K, Kurokawa S, Takamiya A, Fujikawa M, Kinoshita S, Iizuka M, et al. Telepsychiatry versus face-to-face treatment: systematic review and meta-analysis of randomised controlled trials. Br J Psychiatry. 2023;223(3):407-14. doi:10.1192/bjp.2023.86
355. Mohamed A, Elsherif S, Legere B, Fatima N, Shuaib A, Saqqur M. Is telestroke more effective than conventional treatment for acute ischemic stroke? A systematic review and meta-analysis of patient outcomes and thrombolysis rates. Int J Stroke. 2024;19(3):280-92. doi:10.1177/17474930231206066
356. Behrens JR, Kaffes M, Aigner A, Herm J, Erdur H, Siepmann T, et al. Teleneurology vs On-Site Neurology Consultation for Postadmission Hospital Care of Stroke. JAMA Neurol. 2026;83(6):554-62. doi:10.1001/jamaneurol.2026.0615
357. Storjohann S, Kirsch M, Rosenberg B, Rosenberg C, Lange S, Syperek A, et al. The Accuracy of On-Call CT Reporting in Teleradiology Networks in Comparison to In-House Reporting. Healthcare (Basel). 2021;9(4):405. doi:10.3390/healthcare9040405
358. Wittman SR, Hoberman A, Mehrotra A, Sabik LM, Yabes JG, Ray KN. Antibiotic Receipt for Pediatric Telemedicine Visits With Primary Care vs Direct-to-Consumer Vendors. JAMA Netw Open. 2024;7(3):e242359. doi:10.1001/jamanetworkopen.2024.2359
359. Lam K, Lu AD, Shi Y, Covinsky KE. Assessing Telemedicine Unreadiness Among Older Adults in the United States During the COVID-19 Pandemic. JAMA Intern Med. 2020;180(10):1389-91. doi:10.1001/jamainternmed.2020.2671
360. Solimini R, Busardò FP, Gibelli F, Sirignano A, Ricci G. Ethical and Legal Challenges of Telemedicine in the Era of the COVID-19 Pandemic. Medicina (Kaunas). 2021;57(12):1314. doi:10.3390/medicina57121314
361. Elberg JT, Adashi EY. Fraud, Access, and the Future of Telemedicine. J Am Board Fam Med. 2023;36(5):864-6. doi:10.3122/jabfm.2023.230044R1
362. Harris Y, Goldwater JC. Lack of evidence for telehealth fraud. J Telemed Telecare. 2025;31(2):301-5. doi:10.1177/1357633X231177739
363. Ordonanța de urgență a Guvernului nr. 196/2020 pentru modificarea și completarea Legii nr. 95/2006 privind reforma în domeniul sănătății. Monitorul Oficial al României, Partea I. 2020 Nov 19;(1108).
364. Hotărârea Guvernului nr. 1.133/2022 privind aprobarea Normelor metodologice de punere în aplicare a prevederilor Ordonanței de urgență a Guvernului nr. 196/2020 pentru modificarea și completarea Legii nr. 95/2006 privind reforma în domeniul sănătății. Monitorul Oficial al României, Partea I. 2022 Sep 15;(909).
365. Hotărârea Guvernului nr. 521/2023 pentru aprobarea pachetelor de servicii și a Contractului-cadru care reglementează condițiile acordării asistenței medicale, a medicamentelor și a dispozitivelor medicale, în cadrul sistemului de asigurări sociale de sănătate. Monitorul Oficial al României, Partea I. 2023 May 26;(466).
366. Ordinul ministrului sănătății și al președintelui Casei Naționale de Asigurări de Sănătate nr. 1.857/441/2023 privind aprobarea Normelor metodologice de aplicare a Hotărârii Guvernului nr.

521/2023. Monitorul Oficial al României, Partea I. 2023 May 31;(484). Forma consolidată publicată de CNAS, august 2025 [cited 2026 Sep 27]. Available from: https://cnas.ro/wp-content/uploads/2025/08/Norme-1-si-2.pdf

367. Casa Națională de Asigurări de Sănătate. Comunicat: Consultațiile medicale și rețetele electronice pot fi acordate în continuare la distanță [Internet]. București: CNAS; 2023 Jul 6 [cited 2026 Sep 27]. Available from: https://cnas.ro/2023/07/06/comunicat-consultatiile-medicale-si-retetele-electronice-pot-fi-acordate-in-continuare-la-distanta/
368. Ministerul Sănătății. Apeluri PNRR: Componenta 12 Sănătate și Componenta 7 Transformare digitală [Internet]. București: Ministerul Sănătății; 2026 [cited 2026 Sep 27]. Available from: https://www.ms.ro/ro/de-interes/apeluri-pnrr-c12-sanatate/
369. Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients' rights in cross-border healthcare. Official Journal of the European Union. 2011 Apr 4;L 88:45-65.
370. Caffery LJ, Taylor ML, Abbott LM, Janda M, Guitera P, Mar V, et al. Practice Guidelines for Teledermatology in Australia: 2025 Update. Australas J Dermatol. 2025;66(8):e553-64. doi:10.1111/ajd.14599
371. Edwards PJ, Bennett-Britton I, Ridd MJ, Booker M, Barnes RK. Factors affecting the documentation of spoken safety-netting advice in routine GP consultations: a cross-sectional study. Br J Gen Pract. 2021;71(712):e869-76. doi:10.3399/BJGP.2021.0195
372. Medical Device Coordination Group. MDCG 2019-11 rev.1. Qualification and classification of software - Regulation (EU) 2017/745 and Regulation (EU) 2017/746 [Internet]. Brussels: European Commission; 2025 Jun [cited 2026 Sep 27]. Available from: https://health.ec.europa.eu/document/download/b45335c5-1679-4c71-a91c-fc7a4d37f12b_en?filename=mdcg_2019_11_en.pdf
373. Vasiloglou MF, Kotzakioulafi E, Mainardi F, Ahmed M, Staiano AE, Dimidi E, et al. Effectiveness of app-based interventions for glucose management, cardiometabolic and mental health, and related risk factors: an umbrella review with meta-analysis. Lancet Digit Health. 2026;8(8):101009. doi:10.1016/j.landig.2026.101009
374. Sun S, Simonsson O, McGarvey S, Torous J, Goldberg SB. Mobile phone interventions to improve health outcomes among patients with chronic diseases: an umbrella review and evidence synthesis from 34 meta-analyses. Lancet Digit Health. 2024;6(11):e857-70. doi:10.1016/S2589-7500(24)00119-5
375. Moschonis G, Siopis G, Jung J, Eweka E, Willems R, Kwasnicka D, et al. Effectiveness, reach, uptake, and feasibility of digital health interventions for adults with type 2 diabetes: a systematic review and meta-analysis of randomised controlled trials. Lancet Digit Health. 2023;5(3):e125-43. doi:10.1016/S2589-7500(22)00233-3
376. Yap HJ, Lim JJJ, Tan SD, Ang CS. Effectiveness of digital health interventions on adherence and control of hypertension: a systematic review and meta-analysis. J Hypertens. 2024;42(9):1490-504. doi:10.1097/HJH.0000000000003793
377. Chu S, Feng L, Jing H, Zhang D, Li J, Meng X, et al. Efficacy of smartphone apps used alone or with traditional interventions for smoking cessation: a systematic review and meta-analysis. BMJ Evid Based Med. 2026;31(4):242-53. doi:10.1136/bmjebm-2025-113971
378. Linardon J, Torous J, Firth J, Cuijpers P, Messer M, Fuller-Tyszkiewicz M. Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. World Psychiatry. 2024;23(1):139-49. doi:10.1002/wps.21183
379. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the Mobile Application Rating Scale (MARS). PLoS One. 2020;15(11):e0241480. doi:10.1371/journal.pone.0241480
380. Zych MM, Bond R, Mulvenna M, Martinez Carracedo J, Bai L, Leigh S. Quality Assessment of Digital Health Apps: Umbrella Review. J Med Internet Res. 2024;26:e58616. doi:10.2196/58616
381. Jakob R, Harperink S, Rudolf AM, Fleisch E, Haug S, Mair JL, et al. Factors Influencing Adherence to mHealth Apps for Prevention or Management of Noncommunicable Diseases: Systematic Review. J Med Internet Res. 2022;24(5):e35371. doi:10.2196/35371
382. Liu C, Torous J, Fuller-Tyszkiewicz M, Messer M, Anderson C, Soliman OM, et al. Uptake, Adherence, and Attrition in Clinical Trials of Depression and Anxiety Apps: A Systematic Review and Meta-Analysis. JAMA Psychiatry. 2026;83(1):43-50. doi:10.1001/jamapsychiatry.2025.3439
383. Bluetooth Special Interest Group. Generic Health Sensor Profile, version 1.0 [Internet]. Kirkland (WA): Bluetooth SIG; 2023 Jun 13 [cited 2026 Sep 27]. Available from: https://www.bluetooth.com/specifications/specs/generic-health-sensor-profile/

384. HL7 International. Personal Health Device FHIR Implementation Guide, version 2.0.0 (STU 2) [Internet]. Ann Arbor (MI): HL7 International; 2026 Apr 1 [cited 2026 Sep 27]. Available from: https://hl7.org/fhir/uv/phd/
385. Abedian S, Yesakov E, Ostrovskiy S, Hussein R. Integrating Garmin Wearable Data into FHIR-Based Health Systems for Improved Interoperability. Stud Health Technol Inform. 2025;332:185-9. doi:10.3233/SHTI251523
386. Doherty C, Baldwin M, Keogh A, Caulfield B, Argent R. Keeping Pace with Wearables: A Living Umbrella Review of Systematic Reviews Evaluating the Accuracy of Consumer Wearable Technologies in Health Measurement. Sports Med. 2024;54(11):2907-26. doi:10.1007/s40279-024-02077-2
387. Islam SMS, Chow CK, Daryabeygikhotbehsara R, Subedi N, Rawstorn J, Tegegne T, et al. Wearable cuffless blood pressure monitoring devices: a systematic review and meta-analysis. Eur Heart J Digit Health. 2022;3(2):323-37. doi:10.1093/ehjdh/ztac021
388. Rizos EC, Markozannes G, Charitakis N, Filis P, Stoimeni AE, Nørgaard K, et al. Continuous Glucose Monitoring in Type 1 Diabetes, Type 2 Diabetes, and Diabetes During Pregnancy: A Systematic Review with Meta-Analysis of Randomized Controlled Trials. Diabetes Technol Ther. 2025;27(7):537-52. doi:10.1089/dia.2024.0599
389. Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). Häufig gestellte Fragen (FAQ): Änderungen durch das DigiG [Internet]. Bonn: BfArM; 2026 [cited 2026 Sep 27]. Available from: https://www.bfarm.de/DE/Medizinprodukte/_FAQ/DigiG/_artikel.html
390. Kolominsky-Rabas PL, Tauscher M, Gerlach R, Perleth M, Dietzel N. [How robust are studies of currently permanently included digital health applications (DiGA)? Methodological quality of studies demonstrating positive health care effects of DiGA]. Z Evid Fortbild Qual Gesundhwes. 2022;175:1-16. doi:10.1016/j.zefq.2022.09.008
391. Federal Institute for Drugs and Medical Devices (BfArM). Digital Health Applications (DiGA): interesting facts [Internet]. Bonn: BfArM; 2026 [cited 2026 Sep 27]. Available from: https://www.bfarm.de/EN/Medical-devices/Tasks/DiGA-and-DiPA/Digital-Health-Applications/Interesting-facts/_artikel.html
392. Goeldner M, Gehder S. Digital Health Applications (DiGAs) on a Fast Track: Insights From a Data-Driven Analysis of Prescribable Digital Therapeutics in Germany From 2020 to Mid-2024. J Med Internet Res. 2024;26:e59013. doi:10.2196/59013
393. Kawu AA, Hederman L, O'Sullivan D. Exchanging Context and Provenance Based Standardized Patient Generated Health Data with an Electronic Health Record. AMIA Annu Symp Proc. 2024;2024:575-84.
394. Ungurean I, Gherman OI, Lavric A, Dimian M. A privacy-centric microservice framework for secure FHIR-based integration of heterogeneous medical data. Int J Med Inform. 2026;216:106477. doi:10.1016/j.ijmedinf.2026.106477
395. Liu M, Okuhara T, Chang X, Shirabe R, Nishiie Y, Okada H, et al. Performance of ChatGPT Across Different Versions in Medical Licensing Examinations Worldwide: Systematic Review and Meta-Analysis. J Med Internet Res. 2024;26:e60807. doi:10.2196/60807
396. Singhal K, Tu T, Gottweis J, Sayres R, Wulczyn E, Amin M, et al. Toward expert-level medical question answering with large language models. Nat Med. 2025;31(3):943-50. doi:10.1038/s41591-024-03423-7
397. Gong EJ, Bang CS, Lee JJ, Baik GH. Knowledge-Practice Performance Gap in Clinical Large Language Models: Systematic Review of 39 Benchmarks. J Med Internet Res. 2025;27:e84120. doi:10.2196/84120
398. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. JAMA Intern Med. 2023;183(6):589-96. doi:10.1001/jamainternmed.2023.1838
399. Goh E, Gallo R, Hom J, Strong E, Weng Y, Kerman H, et al. Large Language Model Influence on Diagnostic Reasoning: A Randomized Clinical Trial. JAMA Netw Open. 2024;7(10):e2440969. doi:10.1001/jamanetworkopen.2024.40969
400. Goh E, Gallo RJ, Strong E, Weng Y, Kerman H, Freed JA, et al. GPT-4 assistance for improvement of physician performance on patient care tasks: a randomized controlled trial. Nat Med. 2025;31(4):1233-8. doi:10.1038/s41591-024-03456-y
401. Tu T, Schaekermann M, Palepu A, Saab K, Freyberg J, Tanno R, et al. Towards conversational diagnostic artificial intelligence. Nature. 2025;642(8067):442-50. doi:10.1038/s41586-025-08866-7
402. Hager P, Jungmann F, Holland R, Bhagat K, Hubrecht I, Knauer M, et al. Evaluation and mitigation of the limitations of large language models in clinical decision-making. Nat Med. 2024;30(9):2613-22. doi:10.1038/s41591-024-03097-1

403. Bean AM, Payne RE, Parsons G, Kirk HR, Ciro J, Mosquera-Gómez R, et al. Reliability of LLMs as medical assistants for the general public: a randomized preregistered study. Nat Med. 2026;32(2):609-15. doi:10.1038/s41591-025-04074-y
404. Ramaswamy A, Tyagi A, Hugo H, Jiang J, Jayaraman P, Jangda M, et al. ChatGPT Health performance in a structured test of triage recommendations. Nat Med. 2026;32(5):1671-5. doi:10.1038/s41591-026-04297-7
405. Tiller NB, Marcon AR, Zenone M, Kidd KE, Jeukendrup AE, Master Z, et al. Generative artificial intelligence-driven chatbots and medical misinformation: an accuracy, referencing and readability audit. BMJ Open. 2026;16(4):e112695. doi:10.1136/bmjopen-2025-112695
406. Asgari E, Montaña-Brown N, Dubois M, Khalil S, Balloch J, Yeung JA, et al. A framework to assess clinical safety and hallucination rates of LLMs for medical text summarisation. NPJ Digit Med. 2025;8(1):274. doi:10.1038/s41746-025-01670-7
407. Chen S, Gao M, Sasse K, Hartvigsen T, Anthony B, Fan L, et al. When helpfulness backfires: LLMs and the risk of false medical information due to sycophantic behavior. NPJ Digit Med. 2025;8(1):605. doi:10.1038/s41746-025-02008-z
408. Walters WH, Wilder EI. Fabrication and errors in the bibliographic citations generated by ChatGPT. Sci Rep. 2023;13(1):14045. doi:10.1038/s41598-023-41032-5
409. Haltaufderheide J, Ranisch R. The ethics of ChatGPT in medicine and healthcare: a systematic review on Large Language Models (LLMs). NPJ Digit Med. 2024;7(1):183. doi:10.1038/s41746-024-01157-x
410. World Health Organization. WHO releases AI ethics and governance guidance for large multi-modal models [Internet]. Geneva: WHO; 2024 Jan 18 [cited 2026 Sep 27]. Available from: https://www.who.int/news/item/18-01-2024-who-releases-ai-ethics-and-governance-guidance-for-large-multi-modal-models
411. World Health Organization. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models [Internet]. Geneva: WHO; 2025 [cited 2026 Sep 27]. Available from: https://www.who.int/publications/i/item/9789240084759
412. Agenția Națională a Medicamentului și a Dispozitivelor Medicale din România. Dispozitive medicale [Internet]. București: ANMDMR; 2026 [cited 2026 Sep 27]. Available from: https://www.anm.ro/dispozitive-medicale/
413. Winter A, Ammenwerth E, Haux R, Marschollek M, Steiner B, Jahn F. Health Information Systems: Technological and Management Perspectives. 3rd ed. Cham: Springer; 2023. Chapter 2, Basic Concepts and Terms; p. 13-49. doi:10.1007/978-3-031-12310-8_2
414. Winter A, Ammenwerth E, Haux R, Marschollek M, Steiner B, Jahn F. Health Information Systems: Technological and Management Perspectives. 3rd ed. Cham: Springer; 2023. Chapter 4, Management Perspective: Scopes and Tasks of Managing Health Information Systems; p. 153-88. doi:10.1007/978-3-031-12310-8_4
415. Krause T, Jolkver E, Mc Kevitt P, Kramer M, Hemmje M. A Systematic Approach to Diagnostic Laboratory Software Requirements Analysis. Bioengineering (Basel). 2022;9(4):144. doi:10.3390/bioengineering9040144
416. Finnegan H, Mountford N. 25 Years of Electronic Health Record Implementation Processes: Scoping Review. J Med Internet Res. 2025;27:e60077. doi:10.2196/60077
417. Zharima C, Mhlanga S, Abdulla S, Goudge J, Griffiths F. What engagement strategies are useful in facilitating the implementation of electronic health records in health care settings? A rapid review of qualitative evidence synthesis using the normalization process theory. Digit Health. 2024;10:20552076241291286. doi:10.1177/20552076241291286
418. Casa Națională de Asigurări de Sănătate. Comunicat: Aplicațiile informatice ale furnizorilor trebuie actualizate pentru utilizarea noilor funcționalități disponibile prin eSănătateaMea [Internet]. București: CNAS; 2026 Aug 29 [cited 2026 Sep 27]. Available from: https://cnas.ro/2026/08/29/comunicat-aplicatiile-informatice-ale-furnizorilor-trebuie-actualizate-pentru-utilizarea-noilor-functionalitati-disponibile-prin-esanatateamea/
419. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection - Information security management systems - Requirements. 3rd ed. Geneva: International Organization for Standardization; 2022 (amended by ISO/IEC 27001:2022/Amd 1:2024).
420. ISO/IEC 27005:2022. Information security, cybersecurity and privacy protection - Guidance on managing information security risks. 4th ed. Geneva: International Organization for Standardization; 2022.
421. Wasserman L, Wasserman Y. Hospital cybersecurity risks and gaps: Review (for the non-cyber professional). Front Digit Health. 2022;4:862221. doi:10.3389/fdgth.2022.862221

422. Rule A, Chiang MF, Hribar MR. Using electronic health record audit logs to study clinical activity: a systematic review of aims, measures, and methods. J Am Med Inform Assoc. 2020;27(3):480-90. doi:10.1093/jamia/ocz196
423. Sittig DF, Wright A. A guide to mitigating audit log-related risk in medical professional liability cases. J Healthc Risk Manag. 2023;43(2):37-47. doi:10.1002/jhrm.21553
424. Larsen EP, Rao AH, Sasangohar F. Understanding the scope of downtime threats: A scoping review of downtime-focused literature and news media. Health Informatics J. 2020;26(4):2660-72. doi:10.1177/1460458220918539
425. Tully JL, Rao S, Straw I, Gabriel RA, Longhurst CA, Savage S, et al. Patient Care Technology Disruptions Associated With the CrowdStrike Outage. JAMA Netw Open. 2025;8(7):e2530226. doi:10.1001/jamanetworkopen.2025.30226
426. National Institute of Standards and Technology. Advanced Encryption Standard (AES). FIPS PUB 197, updated 2023 May 9. Gaithersburg (MD): NIST; 2023. doi:10.6028/NIST.FIPS.197-upd1
427. National Institute of Standards and Technology. Digital Signature Standard (DSS). FIPS PUB 186-5. Gaithersburg (MD): NIST; 2023 Feb. doi:10.6028/NIST.FIPS.186-5
428. ISO/DIS 17090-4:2025. Health informatics - Public key infrastructure - Part 4: Digital signatures for healthcare documents. Draft International Standard (voting 2025 Jan 27 - 2025 Apr 21). Geneva: International Organization for Standardization; 2025.
429. ISO 17090-1:2021. Health informatics - Public key infrastructure - Part 1: Overview of digital certificate services. Geneva: International Organization for Standardization; 2021.
430. ISO 17090-4:2026. Health informatics - Public key infrastructure - Part 4: Digital signatures for healthcare documents. 3rd ed. Geneva: International Organization for Standardization; 2026.
431. Rescorla E. The Transport Layer Security (TLS) Protocol Version 1.3. RFC 9846 [Internet]. Internet Engineering Task Force; 2026 Jul [cited 2026 Sep 27]. doi:10.17487/RFC9846
432. National Institute of Standards and Technology. Module-Lattice-Based Key-Encapsulation Mechanism Standard. FIPS PUB 203. Gaithersburg (MD): NIST; 2024 Aug. doi:10.6028/NIST.FIPS.203
433. Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC. Official Journal of the European Union. 2014 Aug 28;L 257:73-114.
434. Regulation (EU) 2024/1183 of the European Parliament and of the Council of 11 April 2024 amending Regulation (EU) No 910/2014 as regards establishing the European Digital Identity Framework. Official Journal of the European Union. 2024 Apr 30;L 2024/1183.
435. Ordinul ministrului sănătății și al președintelui CNAS nr. 674/252/2012 privind aprobarea formularului de prescripție medicală electronică pentru medicamente cu și fără contribuție personală în tratamentul ambulatoriu și a Normelor metodologice privind utilizarea și modul de completare a acestuia. Monitorul Oficial al României, Partea I. 2012 Jul 2;(439). Forma consolidată la 29.07.2026.
436. Legea nr. 214/2024 privind utilizarea semnăturii electronice, a mărcii temporale și prestarea serviciilor de încredere bazate pe acestea. Monitorul Oficial al României, Partea I. 2024 Jul 8;(647).
437. European Commission. The legal and technical road to EU Digital Identity Wallets [Internet]. Brussels: European Commission; 2026 [cited 2026 Sep 27]. Available from: https://ec.europa.eu/digital-building-blocks/sites/spaces/EUDIGITALIDENTITYWALLET/pages/915931903/The+legal+and+technical+road+to+EU+Digital+Identity+Wallets
438. International Organization for Standardization. ISO Open Data: ISO deliverables metadata (iso_deliverables_metadata) [Internet]. Geneva: ISO; 2026 [updated 2026 Sep 27; cited 2026 Sep 27]. Licensed under ODC-By 1.0. Available from: https://www.iso.org/open-data.html
439. International Organization for Standardization. ISO/TC 215 Health informatics [Internet]. Geneva: ISO; 2026 [cited 2026 Sep 27]. Available from: https://www.iso.org/committee/54960.html
440. European Committee for Standardization. CEN/TC 251 Health informatics: Business Plan [Internet]. Brussels: CEN; 2022 Dec 13 [cited 2026 Sep 27]. Available from: https://standards.cencenelec.eu/BPCEN/6232.pdf
441. Asociația de Standardizare din România. Buletinul Standardizării, decembrie 2025 [Internet]. București: ASRO; 2025 [cited 2026 Sep 27]. Available from: https://www.asro.ro/wp-content/uploads/2025/12/BS-12-2025.pdf
442. Nifakos S, Chandramouli K, Nikolaou CK, Papachristou P, Koch S, Panaousis E, et al. Influence of Human Factors on Cyber Security within Healthcare Organisations: A Systematic Review. Sensors (Basel). 2021;21(15):5119. doi:10.3390/s21155119

443. Rizzoni F, Magalini S, Casaroli A, Mari P, Dixon M, Coventry L. Phishing simulation exercise in a large hospital: A case study. Digit Health. 2022;8:20552076221081716. doi:10.1177/20552076221081716
444. Yeng PK, Fauzi MA, Vestad A, Yang B, De Moor K, Jacobsen C, et al. Cognitive Dissonance-Based Priming Intervention: Randomized Encouragement With in-the-Wild Phishing Simulation Attack in Health Care. J Med Internet Res. 2026;28:e68051. doi:10.2196/68051
445. Temoshok D, Fenton JL, Choong YY, Lefkovitz N, Regenscheid A, Galluzzo R, et al. Digital Identity Guidelines: Authentication and Lifecycle Management. NIST Special Publication 800-63B-4. Gaithersburg (MD): National Institute of Standards and Technology; 2025. doi:10.6028/NIST.SP.800-63B-4
446. European Union Agency for Cybersecurity (ENISA). Cyber Hygiene in the Health Sector [Internet]. Athens: ENISA; 2025 Sep [cited 2026 Sep 27]. doi:10.2824/3818935
447. Directoratul Național de Securitate Cibernetică. PNRISC - Platforma națională pentru raportarea incidentelor de securitate cibernetică [Internet]. București: DNSC; 2026 [cited 2026 Sep 27]. Available from: https://pnrisc.dnsc.ro/
448. Winter A, Ammenwerth E, Haux R, Marschollek M, Steiner B, Jahn F. Health Information Systems: Technological and Management Perspectives. 3rd ed. Cham: Springer; 2023. doi:10.1007/978-3-031-12310-8
449. Backes J, Geissler A, Subelack J, Ehlig D. Does System Choice Matter? Influence of Hospital Information Systems on Digital Maturity. J Med Syst. 2026;50(1):124. doi:10.1007/s10916-026-02450-w
450. Devin J, Cleary BJ, Cullinan S. The impact of health information technology on prescribing errors in hospitals: a systematic review and behaviour change technique analysis. Syst Rev. 2020;9(1):275. doi:10.1186/s13643-020-01510-7
451. Srinivasamurthy SK, Ashokkumar R, Kodidela S, Howard SC, Samer CF, Chakradhara Rao US. Impact of computerised physician order entry (CPOE) on the incidence of chemotherapy-related medication errors: a systematic review. Eur J Clin Pharmacol. 2021;77(8):1123-31. doi:10.1007/s00228-021-03099-9
452. Taheri Moghadam S, Sadoughi F, Velayati F, Ehsanzadeh SJ, Poursharif S. The effects of clinical decision support system for prescribing medication on patient outcomes and physician practice performance: a systematic review and meta-analysis. BMC Med Inform Decis Mak. 2021;21(1):98. doi:10.1186/s12911-020-01376-8
453. Kalvelage C, Rademacher S, Dohmen S, Marx G, Benstoem C. Decision-Making Authority During Tele-ICU Care Reduces Mortality and Length of Stay-A Systematic Review and Meta-Analysis. Crit Care Med. 2021;49(7):1169-81. doi:10.1097/CCM.0000000000004943
454. Felisberto M, Lima GDS, Celuppi IC, Fantonelli MDS, Zanotto WL, Dias de Oliveira JM, et al. Override rate of drug-drug interaction alerts in clinical decision support systems: A brief systematic review and meta-analysis. Health Informatics J. 2024;30(2):14604582241263242. doi:10.1177/14604582241263242
455. Van Dort BA, Zheng WY, Sundar V, Baysari MT. Optimizing clinical decision support alerts in electronic medical records: a systematic review of reported strategies adopted by hospitals. J Am Med Inform Assoc. 2021;28(1):177-83. doi:10.1093/jamia/ocaa279
456. Chromik J, Klopfenstein SAI, Pfitzner B, Sinno ZC, Arnrich B, Balzer F, et al. Computational approaches to alleviate alarm fatigue in intensive care medicine: A systematic literature review. Front Digit Health. 2022;4:843747. doi:10.3389/fdgth.2022.843747
457. Zhu L, Wei S, An Y, Hu W, Xie X. Severity and Associated Factors of Alarm Fatigue Among Intensive Care Unit Nurses: A Systematic Review and Meta-Analysis. Nurs Crit Care. 2026;31(5):e70666. doi:10.1111/nicc.70666
458. Ministerul Sănătății. Sistem informatic pentru evidența clinică a secțiilor A.T.I. (S.I.E.C.-A.T.I.), cod SMIS 131065 [Internet]. București: Ministerul Sănătății; 2021 [cited 2026 Sep 27]. Available from: https://ms.ro/ro/unitatea-de-implementare-si-coordonare-programe/siec-ati/
459. Toma D, Ghenciu LA, Bedreag OH, Băloi A, Gizea CA, Rițiu SA, et al. Digital Transformation in Critical Care: Implications for Quality of Care, Infection Control, and Clinical Outcomes. J Clin Med. 2025;14(24):8964. doi:10.3390/jcm14248964
460. Vernic C, Ștefănigă SA, Apostol AS, Muntean C. The impact of digitization of intensive care services on patient safety in the Pius Brânzeu Emergency County Hospital in Timișoara [abstract]. Appl Med Inform. 2024;46(Suppl 1):S20. Available from: https://ami.info.umfcluj.ro/index.php/AMI/article/view/1046
461. Vernic C, Bedreag O, Sulițanu DC, Petre I, Muntean LI, Ursoniu S. Digital Hesitancy Among Nurses and Physicians in Anesthesia and Intensive Care: A Cross-Sectional Study from Romania. Nurs Rep. 2026;16(8):290. doi:10.3390/nursrep16080290

462. Lichtner G, Schiefenhövel F, Gashi B, Martin I, Jurth C, Vasiljewa L, et al. Multicenter Evaluation of an Interoperable System for Automated Guideline Adherence Monitoring in ICUs. Crit Care Med. 2026;54(1):119-28. doi:10.1097/CCM.0000000000006961
463. de Andrade JBC, de Medeiros Cavalcante MAO, Lopes TLM, Treigher JMS, Balsells MD, Vasconcelos JL, et al. Discovery of data quality issues in electronic health records: profound consequences for critical care medicine applications - a systematized review. Crit Care. 2026;30(1):19. doi:10.1186/s13054-025-05677-0
464. Vernic C, Tamas TP, Petre I, Ursoniu S. Transforming critical care: the digital revolution's impact on intensive care units. Front Digit Health. 2025;7:1664382. doi:10.3389/fdgth.2025.1664382
465. Institutul Național de Management al Serviciilor de Sănătate. DRG.ro: clasificare DRG, grouper RO DRG v1, noutăți [Internet]. București: INMSS; 2026 [cited 2026 Sep 27]. Available from: https://drg.inmss.ro/
466. Casa Națională de Asigurări de Sănătate. Comunicat: Platforma PIAS va fi înlocuită cu un sistem informatic modern, sigur și accesibil [Internet]. București: CNAS; 2025 Oct 15 [cited 2026 Sep 27]. Available from: https://cnas.ro/2025/10/15/comunicat-platforma-pias-va-fi-inlocuita-cu-un-sistem-informatic-modern-sigur-si-accesibil/
467. Ministerul Sănătății. Componenta 7 - Transformare digitală. Investiția specifică I3.2 - Digitalizarea instituțiilor cu atribuții în domeniul sanitar aflate în subordinea MS [Internet]. București: Ministerul Sănătății; 2023 [cited 2026 Sep 27]. Available from: https://www.ms.ro/ro/de-interes/apeluri-pnrr-c12-sanatate/componenta-7-transformare-digital%C4%83-investitia-specific%C4%83-i32-digitalizarea-institu%C8%9Biilor-cu-atribu%C8%9Bii-%C3%AEn-domeniul-sanitar-aflate-%C3%AEn-subordinea-ms/
468. Regulation (EU) 2021/241 of the European Parliament and of the Council of 12 February 2021 establishing the Recovery and Resilience Facility. Official Journal of the European Union. 2021 Feb 18;L 57:17.
469. Healthcare Information and Management Systems Society (HIMSS). Electronic Medical Record Adoption Model (EMRAM) [Internet]. Chicago: HIMSS; 2026 [cited 2026 Sep 27]. Available from: https://www.himss.org/maturity-models/emram/
470. Snowdon A, Hussein A, Danforth M, Wright A, Oakes R. Digital Maturity as a Predictor of Quality and Safety Outcomes in US Hospitals: Cross-Sectional Observational Study. J Med Internet Res. 2024;26:e56316. doi:10.2196/56316
471. Duncan R, Eden R, Woods L, Wong I, Sullivan C. Synthesizing Dimensions of Digital Maturity in Hospitals: Systematic Review. J Med Internet Res. 2022;24(3):e32994. doi:10.2196/32994
472. Vogel J, Haering A, Kuklinski D, Geissler A. Assessing the Relationship between Hospital Process Digitalization and Hospital Quality - Evidence from Germany. J Med Syst. 2024;48(1):85. doi:10.1007/s10916-024-02101-y
473. Kücking F, Hering A, Strate J, Naumann L, Liebe JD, Hübner UH. Observing the Long Struggle for Digital Transformation in German Hospitals: Results from a Panel Study 2017 to 2025. Stud Health Technol Inform. 2026;340:87-92. doi:10.3233/SHTI260989
474. Eurostat. Hospital discharges and length of stay statistics [Internet]. Luxembourg: Eurostat; 2025 [cited 2026 Sep 27]. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Hospital_discharges_and_length_of_stay_statistics
475. Eurostat. Healthcare resource statistics - beds [Internet]. Luxembourg: Eurostat; 2026 [cited 2026 Sep 27]. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Healthcare_resource_statistics_-_beds
476. DeRienzo C, Levine D, Jain S, Harris A. Postpandemic Recovery of Case Mix Index and Risk-Adjusted Mortality in US Hospitals. JAMA Netw Open. 2025;8(11):e2543398. doi:10.1001/jamanetworkopen.2025.43398
477. Schipmann S, Suero Molina E, Windheuser J, Doods J, Schwake M, Wilbers E, et al. The 30-day readmission rate in neurosurgery-a useful indicator for quality assessment? Acta Neurochir (Wien). 2020;162(11):2659-69. doi:10.1007/s00701-020-04382-x
478. Rabiei R, Almasi S. Requirements and challenges of hospital dashboards: a systematic literature review. BMC Med Inform Decis Mak. 2022;22(1):287. doi:10.1186/s12911-022-02037-8
479. Coiera E, Chan A, Brooke-Cowden K, Rahimi-Ardabili H, Halim N, Tufanaru C. Clinical and economic impact of digital dashboards on hospital inpatient care: a systematic review. JAMIA Open. 2025;8(4):ooaf078. doi:10.1093/jamiaopen/ooaf078
480. Keogh RJ, Harvey H, Brady C, Hassett E, Costelloe SJ, O'Sullivan MJ, et al. Dealing with digital paralysis: Surviving a cyberattack in a National Cancer center. J Cancer Policy. 2024;39:100466. doi:10.1016/j.jcpo.2023.100466

481. Lyon R, Jones A, Burke R, Baysari MT. What Goes Up, Must Come Down: A State-of-the-Art Electronic Health Record Downtime and Uptime Procedure in a Metropolitan Health Setting. Appl Clin Inform. 2023;14(3):513-20. doi:10.1055/s-0043-1768995
482. Aliferis C, Simon G. Artificial Intelligence (AI) and Machine Learning (ML) for Healthcare and Health Sciences: The Need for Best Practices Enabling Trust in AI and ML. In: Simon GJ, Aliferis C, editors. Artificial Intelligence and Machine Learning in Health Care and Medical Sciences: Best Practices and Pitfalls. Cham: Springer; 2024. p. 1-31. doi:10.1007/978-3-031-39355-6_1
483. Aliferis C, Simon G. Foundations and Properties of AI/ML Systems. In: Simon GJ, Aliferis C, editors. Artificial Intelligence and Machine Learning in Health Care and Medical Sciences: Best Practices and Pitfalls. Cham: Springer; 2024. p. 33-94. doi:10.1007/978-3-031-39355-6_2
484. Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Campbell JP. Introduction to Machine Learning, Neural Networks, and Deep Learning. Transl Vis Sci Technol. 2020;9(2):14. doi:10.1167/tvst.9.2.14
485. Simon G, Aliferis C. An Appraisal and Operating Characteristics of Major ML Methods Applicable in Healthcare and Health Science. In: Simon GJ, Aliferis C, editors. Artificial Intelligence and Machine Learning in Health Care and Medical Sciences: Best Practices and Pitfalls. Cham: Springer; 2024. p. 95-195. doi:10.1007/978-3-031-39355-6_3
486. Han R, Acosta JN, Shakeri Z, Ioannidis JPA, Topol EJ, Rajpurkar P. Randomised controlled trials evaluating artificial intelligence in clinical practice: a scoping review. Lancet Digit Health. 2024;6(5):e367-73. doi:10.1016/S2589-7500(24)00047-5
487. Moor M, Banerjee O, Abad ZSH, Krumholz HM, Leskovec J, Topol EJ, et al. Foundation models for generalist medical artificial intelligence. Nature. 2023;616(7956):259-65. doi:10.1038/s41586-023-05881-4
488. Schutte D, Zhang R. Considerations for Specialized Health AI & ML Modelling and Applications: NLP. In: Simon GJ, Aliferis C, editors. Artificial Intelligence and Machine Learning in Health Care and Medical Sciences: Best Practices and Pitfalls. Cham: Springer; 2024. p. 623-41. doi:10.1007/978-3-031-39355-6_14
489. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med. 2023;29(8):1930-40. doi:10.1038/s41591-023-02448-8
490. Aliferis C, Simon G. Lessons Learned from Historical Failures, Limitations and Successes of AI/ML in Healthcare and the Health Sciences. Enduring Problems, and the Role of Best Practices. In: Simon GJ, Aliferis C, editors. Artificial Intelligence and Machine Learning in Health Care and Medical Sciences: Best Practices and Pitfalls. Cham: Springer; 2024. p. 543-606. doi:10.1007/978-3-031-39355-6_12
491. Aliferis C, Simon G. The Development Process and Lifecycle of Clinical Grade and Other Safety and Performance-Sensitive AI/ML Models. In: Simon GJ, Aliferis C, editors. Artificial Intelligence and Machine Learning in Health Care and Medical Sciences: Best Practices and Pitfalls. Cham: Springer; 2024. p. 289-340. doi:10.1007/978-3-031-39355-6_6
492. Collins GS, Dhiman P, Ma J, Schlussel MM, Archer L, Van Calster B, et al. Evaluation of clinical prediction models (part 1): from development to external validation. BMJ. 2024;384:e074819. doi:10.1136/bmj-2023-074819
493. Tejani AS, Klontzas ME, Gatti AA, Mongan JT, Moy L, Park SH, et al. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 Update. Radiol Artif Intell. 2024;6(4):e240300. doi:10.1148/ryai.240300
494. Chen RJ, Wang JJ, Williamson DFK, Chen TY, Lipkova J, Lu MY, et al. Algorithmic fairness in artificial intelligence for medicine and healthcare. Nat Biomed Eng. 2023;7(6):719-42. doi:10.1038/s41551-023-01056-8
495. Van Calster B, Steyerberg EW, Wynants L, van Smeden M. There is no such thing as a validated prediction model. BMC Med. 2023;21(1):70. doi:10.1186/s12916-023-02779-w
496. Finlayson SG, Subbaswamy A, Singh K, Bowers J, Kupke A, Zittrain J, et al. The Clinician and Dataset Shift in Artificial Intelligence. N Engl J Med. 2021;385(3):283-6. doi:10.1056/NEJMc2104626
497. National Academy of Medicine; Krishnan S, Elliott A, editors. Generative Artificial Intelligence in Health and Medicine: Opportunities and Responsibilities for Transformative Innovation. Washington (DC): National Academies Press; 2025. doi:10.17226/28907
498. Kapoor S, Narayanan A. Leakage and the reproducibility crisis in machine-learning-based science. Patterns (N Y). 2023;4(9):100804. doi:10.1016/j.patter.2023.100804
499. Monaghan TF, Rahman SN, Agudelo CW, Wein AJ, Lazar JM, Everaert K, et al. Foundational Statistical Principles in Medical Research: Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value. Medicina (Kaunas). 2021;57(5):503. doi:10.3390/medicina57050503

500. Van Calster B, Collins GS, Vickers AJ, Wynants L, Kerr KF, Barreñada L, et al. Evaluation of performance measures in predictive artificial intelligence models to support medical decisions: overview and guidance. Lancet Digit Health. 2025;7(12):100916. doi:10.1016/j.landig.2025.100916
501. Çorbacıoğlu ŞK, Aksel G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: A guide to interpreting the area under the curve value. Turk J Emerg Med. 2023;23(4):195-8. doi:10.4103/tjem.tjem_182_23
502. Wong A, Otles E, Donnelly JP, Krumm A, McCullough J, DeTroyer-Cooley O, et al. External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients. JAMA Intern Med. 2021;181(8):1065-70. doi:10.1001/jamainternmed.2021.2626
503. Wong A, Currey D, Schwinne M, Park-Egan B, Meyer S, Gutting A, et al. Multicenter Prospective Validation of an Updated Proprietary Sepsis Prediction Model. JAMA Netw Open. 2026;9(2):e260181. doi:10.1001/jamanetworkopen.2026.0181
504. Vickers AJ, Holland F. Decision curve analysis to evaluate the clinical benefit of prediction models. Spine J. 2021;21(10):1643-8. doi:10.1016/j.spinee.2021.02.024
505. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378. doi:10.1136/bmj-2023-078378
506. Moons KGM, Damen JAA, Kaul T, Hooft L, Andaur Navarro C, Dhiman P, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ. 2025;388:e082505. doi:10.1136/bmj-2024-082505
507. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364-74. doi:10.1038/s41591-020-1034-x
508. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, SPIRIT-AI and CONSORT-AI Working Group, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Nat Med. 2020;26(9):1351-63. doi:10.1038/s41591-020-1037-7
509. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33. doi:10.1038/s41591-022-01772-9
510. Sounderajah V, Guni A, Liu X, Collins GS, Karthikesalingam A, Markar SR, et al. The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence. Nat Med. 2025;31(10):3283-9. doi:10.1038/s41591-025-03953-8
511. Gallifant J, Afshar M, Ameen S, Aphinyanaphongs Y, Chen S, Cacciamani G, et al. The TRIPOD-LLM reporting guideline for studies using large language models. Nat Med. 2025;31(1):60-9. doi:10.1038/s41591-024-03425-5
512. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med. 2022;28(1):31-8. doi:10.1038/s41591-021-01614-0
513. Plana D, Shung DL, Grimshaw AA, Saraf A, Sung JJY, Kann BH. Randomized Clinical Trials of Machine Learning Interventions in Health Care: A Systematic Review. JAMA Netw Open. 2022;5(9):e2233946. doi:10.1001/jamanetworkopen.2022.33946
514. Makar J, Abdelmalak J, Con D, Hafeez B, Garg M. Use of artificial intelligence improves colonoscopy performance in adenoma detection: a systematic review and meta-analysis. Gastrointest Endosc. 2025;101(1):68-81.e8. doi:10.1016/j.gie.2024.08.033
515. Lukac PJ, Turner W, Vangala S, Chin AT, Khalili J, Shih YT, et al. Ambient AI Scribes in Clinical Practice: A Randomized Trial. NEJM AI. 2025;2(12). doi:10.1056/aioa2501000
516. Afshar M, Baumann MR, Resnik F, Hintzke J, Sullivan AG, Wills G, et al. A Pragmatic Randomized Controlled Trial of Ambient Artificial Intelligence to Improve Health Practitioner Well-Being. NEJM AI. 2025;2(12). doi:10.1056/aioa2500945
517. Chowdhury A, Casey M, Wilson J, Pollak KI, Goldstein BA, Bedoya A, et al. Comparing ambient scribes: a randomized crossover clinical trial addressing ambient scribe technologies' impact on physician burnout. J Am Med Inform Assoc. 2026;33(5):990-9. doi:10.1093/jamia/ocag018
518. Arun N, Gaw N, Singh P, Chang K, Aggarwal M, Chen B, et al. Assessing the Trustworthiness of Saliency Maps for Localizing Abnormalities in Medical Imaging. Radiol Artif Intell. 2021;3(6):e200267. doi:10.1148/ryai.2021200267
519. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health. 2021;3(11):e745-50. doi:10.1016/S2589-7500(21)00208-9

520. Gichoya JW, Banerjee I, Bhimireddy AR, Burns JL, Celi LA, Chen LC, et al. AI recognition of patient race in medical imaging: a modelling study. Lancet Digit Health. 2022;4(6):e406-14. doi:10.1016/S2589-7500(22)00063-2
521. National Academy of Medicine; Adams L, Fontaine E, Matheny M, Krishnan S, editors. An Artificial Intelligence Code of Conduct for Health and Medicine: Essential Guidance for Aligned Action. Washington (DC): National Academies Press; 2025. doi:10.17226/29087
522. Alber DA, Yang Z, Alyakin A, Yang E, Rai S, Valliani AA, et al. Medical large language models are vulnerable to data-poisoning attacks. Nat Med. 2025;31(2):618-26. doi:10.1038/s41591-024-03445-1
523. European Commission. AI Act: regulatory framework on artificial intelligence [Internet]. Brussels: European Commission; 2026 [cited 2026 Sep 27]. Available from: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
524. European Commission. Proposal for a Regulation of the European Parliament and of the Council amending Regulations (EU) 2017/745 and (EU) 2017/746 as regards simplifying and reducing the burden of the rules on medical devices and in vitro diagnostic medical devices, and amending Regulation (EU) 2022/123 and Regulation (EU) 2024/1689. COM(2025) 1023 final, 2025/0404 (COD) [Internet]. Strasbourg: European Commission; 2025 Dec 16 [cited 2026 Sep 27]. Available from: https://health.ec.europa.eu/document/download/25e7ea7c-cab3-40cf-86d9-d11f5e7744d8_en?filename=md_com_2025-1023_act_en.pdf
525. Ştefănigă SA, Cordoş AA, Ivascu T, Feier CVI, Muntean C, Stupinean CV, et al. Advancing Precision Oncology with Digital and Virtual Twins: A Scoping Review. Cancers (Basel). 2024;16(22):3817. doi:10.3390/cancers16223817
526. Council of the European Union. Artificial intelligence: Council and Parliament agree to simplify and streamline rules [Internet]. Brussels: Council of the EU; 2026 May 7 [updated 2026 May 18; cited 2026 Sep 27]. Available from: https://www.consilium.europa.eu/en/press/press-releases/2026/05/07/artificial-intelligence-council-and-parliament-agree-to-simplify-and-streamline-rules/
527. Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act). Official Journal of the European Union. 2023 Dec 22;L 2023/2854.
528. European Parliament. Legislative Train Schedule: The Digital Omnibus Regulation proposal, 2025/0360(COD) [Internet]. Brussels: European Parliament; 2026 Aug 1 [cited 2026 Sep 27]. Available from: https://www.europarl.europa.eu/legislative-train/theme-a-new-plan-for-europe-s-sustainable-prosperity-and-competitiveness/file-digital-package
529. Regenstrief Institute. Download LOINC: LOINC version 2.83 [Internet]. Indianapolis (IN): Regenstrief Institute; 2026 Aug 19 [cited 2026 Sep 27]. Available from: https://loinc.org/downloads/
530. Google. Share files from Google Drive [Internet]. Google Drive Help; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/drive/answer/2494822
531. Google. Find what's changed in a file [Internet]. Google Docs Editors Help; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/docs/answer/190843
532. Google. Create a strong password & a more secure account [Internet]. Google Account Help; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/accounts/answer/32040
533. Google. Turn on 2-Step Verification [Internet]. Google Account Help; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/accounts/answer/185839
534. Cybersecurity and Infrastructure Security Agency. #StopRansomware Guide [Internet]. Washington (DC): CISA; 2023 Sep [cited 2026 Sep 27]. Available from: https://www.cisa.gov/stopransomware/ransomware-guide
535. National Library of Medicine (US). MeSH Browser [Internet]. Bethesda (MD): NLM; 2026 [cited 2026 Sep 27]. Available from: https://meshb.nlm.nih.gov/
536. Schiavenato M, Chu F. PICO: What it is and what it is not. Nurse Educ Pract. 2021;56:103194. doi:10.1016/j.nepr.2021.103194
537. Farič N, Potts HW, Heilman JM. Quality of Male and Female Medical Content on English-Language Wikipedia: Quantitative Content Analysis. J Med Internet Res. 2024;26:e47562. doi:10.2196/47562
538. Marquez-Romero JM, Lee Á, Soto-Cabrera E, Hernández-Curiel BC, Prado-Aguilar CA. Content and Quality of Information about Stroke in Wikipedia across Multiple Languages. Eur Neurol. 2022;85(4):308-12. doi:10.1159/000521938
539. Halboub E, Al-Ak'hali MS, Alqahtani AS, Abdulghani EA, Kamil MA, Alhajj MN, et al. Quality of web-based Arabic health information on dental implants: an infodemiological study. BMC Oral Health. 2023;23(1):232. doi:10.1186/s12903-023-02938-8

540. Cherrez-Ojeda I, Zuberbier T, Rodas-Valero G, Sanchez J, Rudenko M, Dramburg S, et al. Evaluation of the Quality and Reliability of ChatGPT-4's Responses on Allergen Immunotherapy Using Validated Instruments for Health Information Quality Assessment. Clin Transl Allergy. 2025;15(12):e70130. doi:10.1002/clt2.70130
541. Gerek ME, Önalan T, Çölkesen F, Arslan Ş. Evaluating large language models for WAO/EAACI guideline compliance in hereditary angioedema management. Allergol Immunopathol (Madr). 2025;53(4):51-9. doi:10.15586/aei.v53i4.1353
542. CHART Collaborative. Reporting guideline for chatbot health advice studies: the Chatbot Assessment Reporting Tool (CHART) statement. BMJ Med. 2025;4(1):e001632. doi:10.1136/bmjmed-2025-001632
543. World Health Organization. Hypertension. Fact sheet [Internet]. Geneva: WHO; 2025 Sep 25 [cited 2026 Sep 27]. Available from: https://www.who.int/news-room/fact-sheets/detail/hypertension
544. Google. Create and use pivot tables. Google Docs Editors Help [Internet]. Mountain View (CA): Google; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/docs/answer/1272900
545. World Health Organization. ICD-11 Coding Tool, Mortality and Morbidity Statistics (MMS), release 2026-01 [Internet]. Geneva: WHO; 2026 [cited 2026 Sep 27]. Available from: https://icd.who.int/ct
546. World Health Organization. ICD API [Internet]. Geneva: WHO; 2026 [cited 2026 Sep 27]. Available from: https://icd.who.int/icdapi
547. Digital Scholar. Zotero 10: download [Internet]. Digital Scholar; c2026 [cited 2026 Sep 27]. Available from: https://www.zotero.org/download/
548. Digital Scholar. Zotero documentation: Word processor plugin integration [Internet]. Digital Scholar; c2026 [cited 2026 Sep 27]. Available from: https://www.zotero.org/support/word_processor_integration
549. Digital Scholar. Zotero documentation: Citation styles [Internet]. Digital Scholar; c2026 [cited 2026 Sep 27]. Available from: https://www.zotero.org/support/styles
550. Elsevier. Mendeley Reference Manager for Windows, version 2.149.0 [Internet]. Elsevier; c2026 [cited 2026 Sep 27]. Available from: https://www.mendeley.com/download-reference-manager/windows
551. Digital Scholar. Zotero documentation: How do I export my Zotero library? [Internet]. Digital Scholar; [updated 2025 Nov 24; cited 2026 Sep 27]. Available from: https://www.zotero.org/support/kb/exporting
552. Elsevier. How can I add new references to my Library in Mendeley Reference Manager? [Internet]. Elsevier Support Center; c2026 [cited 2026 Sep 27]. Available from: https://www.elsevier.support/mendeley/answer/how-can-i-add-new-references-to-my-library-in-mendeley-reference-manager
553. PhysioNet. EEG Motor Movement/Imagery Dataset, version 1.0.0 [Internet]. Cambridge (MA): MIT Laboratory for Computational Physiology; 2026 [cited 2026 Sep 27]. Available from: https://physionet.org/content/eegmmidb/1.0.0/
554. MIT Laboratory for Computational Physiology. wfdb: the WFDB Python package, version 4.3.1 [Internet]. Python Package Index; 2026 [cited 2026 Sep 27]. Available from: https://pypi.org/project/wfdb/
555. WFDB Python package documentation, version 4.3.1: WFDB records and annotations (rdrecord, rdann) [Internet]. Read the Docs; 2026 [cited 2026 Sep 27]. Available from: https://wfdb.readthedocs.io/en/latest/io.html
556. Schroeder AB, Dobson ETA, Rueden CT, Tomancak P, Jug F, Eliceiri KW. The ImageJ ecosystem: Open-source software for image visualization, processing, and analysis. Protein Sci. 2021;30(1):234-49. doi:10.1002/pro.3993
557. ImageJ community. Fiji Downloads [Internet]. imagej.net; 2026 [cited 2026 Sep 27]. Available from: https://imagej.net/software/fiji/downloads
558. Rasband W. ImageJ Release Notes, version 1.54t [Internet]. Bethesda (MD): National Institutes of Health; 2026 May 16 [cited 2026 Sep 27]. Available from: https://wsr.imagej.net/notes.html
559. ImageJ. ImageJ Documentation: Menu commands [Internet]. Bethesda (MD): National Institutes of Health; [cited 2026 Sep 27]. Available from: https://imagej.net/ij/docs/menus/file.html
560. The Cancer Imaging Archive. Data Usage Policies and Restrictions [Internet]. TCIA; [cited 2026 Sep 27]. Available from: https://www.cancerimagingarchive.net/data-usage-policies-and-restrictions/
561. Gahm NA, Rueden CT, Evans EL, Selzer G, Hiner MC, Chacko JV, et al. New Extensibility and Scripting Tools in the ImageJ Ecosystem. Curr Protoc. 2021;1(8):e204. doi:10.1002/cpz1.204
562. Rueden CT, Hiner MC, Evans EL, Pinkert MA, Lucas AM, Carpenter AE, et al. PyImageJ: A library for integrating ImageJ and Python. Nat Methods. 2022;19(11):1326-7. doi:10.1038/s41592-022-01655-4
563. OpenEMR Project. OpenEMR 8.4.1 release [Internet]. GitHub; 2026 Sep 20 [cited 2026 Sep 27]. Available from: https://github.com/openemr/openemr/releases

564. OpenEMR Project. OpenEMR online demo (version 8.4.1) [Internet]. OpenEMR Foundation; 2026 [cited 2026 Sep 27]. Available from: https://www.open-emr.org/demo/
565. OpenEMR Project. OpenEMR FHIR API (FHIR_README) [Internet]. GitHub; 2026 [cited 2026 Sep 27]. Available from: https://github.com/openemr/openemr/blob/master/FHIR_README.md
566. HAPI FHIR Project. HAPI FHIR Test/Demo Server, R4 endpoint [Internet]. 2026 [cited 2026 Sep 27]. Available from: https://hapi.fhir.org/baseR4
567. HL7 International. FHIR R5 (v5.0.0). Observation Vital Signs Profile [Internet]. Ann Arbor (MI): HL7 International; 2023 [cited 2026 Sep 27]. Available from: https://hl7.org/fhir/R5/observation-vitalsigns.html
568. OpenEMR Project. HOWTO: Create a New Patient Record - OpenEMR v7 [Internet]. OpenEMR Project Wiki; 2026 [cited 2026 Sep 27]. Available from: https://www.open-emr.org/wiki/index.php/HOWTO:_Create_a_New_Patient_Record_-_OpenEMR_v7
569. Murphy EM, Stein A, Pahwa R, McGuire M, Kumra T. Difference in medical student performance in a standardized patient encounter between telemedicine and in-person environments. Med Educ Online. 2024;29(1):2388422. doi:10.1080/10872981.2024.2388422
570. Gunner CK, Eisner E, Watson AJ, Duncan JL. Teaching webside manner: development and initial evaluation of a video consultation skills training module for undergraduate medical students. Med Educ Online. 2021;26(1):1954492. doi:10.1080/10872981.2021.1954492
571. Jitsi. Jitsi Meet, stable release 2.0.11248 [software on the Internet]. 2026 Sep 14 [cited 2026 Sep 27]. Available from: https://github.com/jitsi/jitsi-meet/releases
572. Pozo P, Landino MC, Maga JM, Jorge LM, Araya R, Horn DB, et al. Situational Awareness in Telehealth: A Virtual Standardized Patient Case for Transitioning Preclinical to Clinical Medical Students. MedEdPORTAL. 2025;21:11517. doi:10.15766/mep_2374-8265.11517
573. Lesselroth B, Monkman H, Palmer R, Kuziemsky C, Liew A, Foulks K, et al. Assessing Telemedicine Competencies: Developing and Validating Learner Measures for Simulation-Based Telemedicine Training. AMIA Annu Symp Proc. 2023;2023:474-83.
574. Hart A, Romney D, Sarin R, Mechanic O, Hertelendy AJ, Larson D, et al. Developing Telemedicine Curriculum Competencies for Graduate Medical Education: Outcomes of a Modified Delphi Process. Acad Med. 2022;97(4):577-85. doi:10.1097/ACM.0000000000004463
575. World Health Organization. WHO guidelines [Internet]. Geneva: WHO; 2026 [cited 2026 Sep 27]. Available from: https://www.who.int/publications/who-guidelines
576. Adobe. Create self-signed digital IDs in Acrobat [Internet]. Adobe Help Center; 2025 Sep 23 [cited 2026 Sep 27]. Available from: https://helpx.adobe.com/acrobat/desktop/protect-documents/manage-digital-ids/create-self-signed-id.html
577. Adobe. Validate digital signatures in Adobe Acrobat [Internet]. Adobe Help Center; 2025 Sep 23 [cited 2026 Sep 27]. Available from: https://helpx.adobe.com/acrobat/desktop/e-sign-documents/manage-digital-signatures/validate-digital-sign.html
578. Google. Check if a site's connection is secure [Internet]. Google Chrome Help; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/chrome/answer/95617
579. European Commission. EU/EEA Trusted List Browser, eIDAS Dashboard [Internet]. Brussels: European Commission; 2026 [cited 2026 Sep 27]. Available from: https://eidas.ec.europa.eu/efda/trust-services/browse/eidas/tls
580. JGraph Ltd. draw.io (diagrams.net) [Internet]. Northampton: JGraph; 2026 [cited 2026 Sep 27]. Available from: https://www.drawio.com/
581. JGraph Ltd. drawio-desktop v31.5.3 [software]. GitHub; 2026 Sep 26 [cited 2026 Sep 27]. Available from: https://github.com/jgraph/drawio-desktop/releases/tag/v31.5.3
582. JGraph Ltd. Flowcharts with swimlanes [Internet]. draw.io documentation; 2026 [cited 2026 Sep 27]. Available from: https://www.drawio.com/docs/diagram-types/swimlane-diagrams/
583. Google. Google Sheets function list [Internet]. Mountain View (CA): Google; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/docs/table/25273?hl=en
584. Microsoft. SUMPRODUCT function [Internet]. Redmond (WA): Microsoft; 2026 [cited 2026 Sep 27]. Available from: https://support.microsoft.com/en-us/office/sumproduct-function-16753e75-9f68-4874-94ac-4d2145a2fd2e
585. Antonacci G, Lennox L, Barlow J, Evans L, Reed J. Process mapping in healthcare: a systematic review. BMC Health Serv Res. 2021;21(1):342. doi:10.1186/s12913-021-06254-1

586. Gartner JB, Landa P, Haren MT, Lemaire C, Tanfani E, Paquet C, et al. The Implementation of a Business Process Model and Notation for Modeling Patient Health Care Trajectories: Systematic Review. J Med Internet Res. 2026;28:e78506. doi:10.2196/78506

587. Google. Add & edit a chart or graph. Google Docs Editors Help [Internet]. Mountain View (CA): Google; 2026 [cited 2026 Sep 27]. Available from: https://support.google.com/docs/answer/63824

588. The Document Foundation. Download LibreOffice (version 26.8.0) [Internet]. Berlin: The Document Foundation; 2026 [cited 2026 Sep 27]. Available from: https://www.libreoffice.org/download/download-libreoffice/