



**HAL**  
open science

# Dialogue humain machine pour l'aide à la décision médicale

Chadia Ed-Driouch

► **To cite this version:**

Chadia Ed-Driouch. Dialogue humain machine pour l'aide à la décision médicale. Intelligence artificielle [cs.AI]. Nantes Université, 2023. Français. NNT : 2023NANU4027 . tel-04334708

**HAL Id: tel-04334708**

**<https://theses.hal.science/tel-04334708v1>**

Submitted on 11 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THESE DE DOCTORAT

NANTES UNIVERSITE

ECOLE DOCTORALE N° 603

Education, Cognition, Langages, Interactions, Santé

Spécialité : Informatique

Par

**Chadia ED-DRIOUCH**

**Dialogue humain machine pour l'aide à la décision médicale**

Thèse présentée et soutenue à Nantes, le 22 Mai 2023 Unité de recherche :

Laboratoire des Sciences du Numérique de Nantes (LS2N)

## Rapporteuses avant soutenance :

Céline LOUAPRE Maîtresse de conférences - Praticienne hospitalière HDR, Institut du Cerveau, Sorbonne Université

Isabelle PECCI Maîtresse de conférences HDR, Université de Lorraine

## Composition du Jury :

Président : Gilles EDAN

Professeur d'université - Praticien hospitalier, CHU de Rennes

Dir. de thèse : Franck MARS

Directeur de recherche, CNRS, Nantes Université

Co-dir. de thèse : Pierre-Antoine GOURRAUD

Professeur d'université – Praticien hospitalier, CR2TI,  
Nantes Université

Encadrant de thèse : Cédric DUMAS

Maître de conférences, IMT-Atlantique

## Invité

Pascal SEMPÉ VP / Directeur Senior R&D, Medexprim



*« Dès que tu avances sur le chemin, le chemin apparaît. »*

Djalâl ad-Dîn Rûmi





# RÉSUMÉ

Les progrès en science des données, informatique et biotechnologie, avec la disponibilité de vastes quantités de données sur les patients et les maladies, font que cette époque est plus propice que jamais à l'avancement de la recherche translationnelle, en particulier pour la prise en charge des maladies complexes. Ces maladies, souvent multifactorielles par nature, nécessitent une prise en charge individualisée pour offrir une vision plus précise de la pathologie et ainsi améliorer la qualité des soins apportés aux patients. L'individualisation des soins constitue depuis toujours l'axe principal du développement des approches de la Médecine de Précision (MP), avec pour objectif la délivrance de mesures spécifiques permettant de proposer un protocole de prise en charge spécialement adapté à chaque patient, selon des critères propres à sa pathologie. Plusieurs solutions ont émergé de ce concept afin d'apporter un appui véritable pour accélérer l'identification des maladies et aider à leur gestion. Cependant, l'acceptabilité et l'adoption de ces nouvelles pratiques du numérique dans les services de santé, sont limitées par des enjeux liés aux données (utilité clinique, qualité, confidentialité, transparence, etc.) et aux méthodes (performance, clarté des processus sous-jacents, etc.) qui les sous-tendent.

La collaboration accompagnée d'un dialogue clair et constructif est un des éléments clés pour se comprendre et connaître les préoccupations de l'utilisateur. Établie comme principe de conception, la collaboration humain-algorithme permet de mieux orienter et développer des solutions appropriées. Dans cette optique, cette thèse s'intéresse aux mécanismes du dialogue humain-machine, à travers les interactions avec les données et les algorithmes, et leurs visualisations.

Notre premier objectif concerne l'utilisation des données de patients atteints de Sclérose En Plaques (SEP) pour développer un prototype fonctionnel de système d'aide à la décision clinique (CDSS - *Clinical Decision Support System*), basé sur les concepts de la MP. Grâce à une collaboration avec des neurologues, cette thèse a abouti au développement d'un CDSS, nommé MS Vista. MS Vista est basé sur une conception centrée sur l'utilisateur et offre une interface humain-machine ergonomique permettant d'interagir avec les données. Il est fondé sur la distinction entre le POI (*Patient Of Interest*), le patient pour lequel une décision médicale doit être prise, et les PORs (*Patients Of Reference*), des patients

présentant des facteurs pertinents de la maladie, similaires à ceux du POI, validées dans la littérature et par des neurologues experts. Il exploite la qualité des données d'un essai clinique de phase III (NCT00906399) pour contextualiser le POI et faire des projections de données de ses PORs. Le but est d'aider les cliniciens de la SEP à établir un pronostic personnalisé. Par ailleurs, il offre au médecin un contrôle direct et lui permet d'examiner la qualité et la quantité de données des PORs sélectionnés pour chaque POI. Ainsi, il peut déterminer si les données sont favorables à la prise de décision. MS Vista a également le potentiel d'enrichir la communication médecin-patient grâce à son mode de visualisation dual et interactif. Il montre comment les différents facteurs de la maladie contribuent à sa progression, tout en illustrant les avantages potentiels de chaque option thérapeutique, pour aider à choisir le traitement le plus approprié pour le POI. L'évaluation du CDSS auprès des neurologues a révélé leur appréciation quant à la capacité de MS Vista à contextualiser le patient et à illustrer la probabilité raisonnable de demeurer dans un état donné selon différents scénarios thérapeutiques. Ils ont également indiqué qu'il pouvait aider à la prise de décision et à la communication avec le patient.

Le second objectif de cette thèse est de concevoir des modèles de l'apprentissage automatique (ML - *Machine Learning*) qui collaborent avec les médecins afin qu'ils soient pertinents et compréhensibles. Le but est d'aider à relever les défis liés à l'adoption de systèmes basés sur l'intelligence artificielle dans le domaine des soins de santé. Nous avons proposé une nouvelle approche d'intelligence hybride tirant parti des intelligences humaine et artificielle, pour développer des modèles prédictifs basés sur la pratique médicale à travers la collaboration médecin-algorithme. Cette démarche pourrait améliorer la qualité des résultats d'une part et renforcer la confiance des médecins d'autre part, en étant rassurés par leur implication dans le processus de prédiction. À cette fin, nous avons exploité la qualité des données d'un essai clinique, la capacité des algorithmes de ML et le raisonnement des neurologues experts pour créer des modèles hybrides de prédiction de la progression de la SEP. Au lieu d'entraîner les algorithmes de ML sur les données brutes pour créer des modèles prédictifs (méthode conventionnelle), nous proposons de pré-catégoriser les données sur la base des connaissances et de l'expérience des médecins experts. La comparaison des résultats de la méthode conventionnelle de ML avec la méthode hybride médecin-algorithme a montré que les modèles basés sur cette dernière sont capables de donner de meilleurs résultats. Cette approche de l'intelligence artificielle col-

laborative ouvre des perspectives prometteuses sur la collaboration humains algorithmes pour fournir des solutions cohérentes d'aide à la prise de décision personnalisée.

Les travaux présentés dans cette thèse illustrent la complémentarité de l'humain et du numérique pour un avenir plus sûr de la recherche translationnelle.

**Mot clés :** Interaction Humain-Machine, Intelligence Hybride, Intelligence Artificielle, Apprentissage Automatique, Système d'Aide à la Décision Clinique, Collaboration Humain-Algorithmes, Sclérose En Plaques.



# ABSTRACT

Advances in data science, computer science and biotechnology, along with the availability of vast amounts of patient and disease data, make this time more opportune than ever for the advancement of translational research, especially for the management of complex diseases. Such diseases, often multifactorial by nature, require individualized management to provide a more precise view of the pathology and thus improve the quality of patient care. The individualization of care has always been the focus of the development of Precision Medicine (PM) approaches, with the purpose of delivering specific measures enabling to propose a tailored management protocol for each patient according to criteria specific to their pathology. Several solutions have emerged from this concept in order to provide real support to speed up the identification of diseases and help in their management. However, issues related to the underlying data (clinical utility, quality, confidentiality, transparency, etc.) and methods (performance, clarity of underlying processes, etc.) limit the acceptability and adoption of these new digital practices in health services.

Collaboration with a clear and constructive dialogue is one of the key elements to understand each other and to know the user's concerns. As a design principle, human-algorithm collaboration allows to better drive and develop appropriate solutions. In this perspective, this thesis focuses on the mechanisms of human-computer dialogue, through interactions with data and algorithms, and their visualizations.

Our first purpose in this thesis is to use data from Multiple Sclerosis (MS) patients to develop a functional prototype of a Clinical Decision Support System (CDSS) based on PM concepts. Thanks to a collaboration with neurologists, this thesis has resulted the development of a CDSS, called MS Vista. MS Vista is based on a user-centered design and provides an ergonomic human-computer interface to interact with the data. It is built on the distinction between the POI (Patient Of Interest), the patient for whom a medical decision has to be made and the PORs (Patients Of Reference), patients with relevant disease features, similar to those of the POI, validated in the literature and by expert neurologists. It leverages the data quality of a Phase III clinical trial (NCT00906399) to contextualize the POI and make data projections of its PORs. The aim is to help MS clinicians make a personalized prognosis. In addition, it provides the physician with direct control and

allows them to examine the quality and quantity of the selected PORs data for each POI. This enables the physician to determine whether the data is suitable for decision-making. MS Vista has also the potential to enrich physician-patient communication through its dual and interactive visualization mode. It illustrates how different features of the disease contribute to its progression, as well as the potential benefits of each treatment option, to assist in choosing the most appropriate treatment for the POI. Evaluation of the CDSS with neurologists revealed their appreciation for MS Vista's ability to contextualize the patient and illustrate the reasonable likelihood of remaining in a given condition under different therapeutic scenarios. They also reported that it could assist in decision-making and communication with the patient.

The second objective of this thesis is to conceive Machine Learning (ML) models that collaborate with physicians to be relevant and understandable. The aim is to help address the challenges associated with the adoption of artificial intelligence-based systems in healthcare. We proposed a new hybrid intelligence approach leveraging human and artificial intelligences to develop predictive models based on medical practice through physician-algorithm collaboration. This may enhance the quality of the results first, and second, enhance physician confidence through reassurance of their inclusion in the prediction process. To this end, we exploited the data quality of a clinical trial, the capability of ML algorithms, and the reasoning of expert neurologists to create hybrid models for predicting MS progression. Instead of training ML algorithms on the raw data to create predictive models (conventional method), we propose to pre-categorize data based on expert physicians' knowledge and experience. The results of the conventional ML method compared with the hybrid physician-algorithm method showed that the models based on the latter are able to perform better. This approach of collaborative artificial intelligence opens promising perspectives on human-algorithm collaboration to provide coherent solutions for personalized decision support.

The works of this thesis illustrate the complementarity of human and digital technology for a more secure future of translational research.

**Keywords :** Human-Computer Interaction, Hybrid Intelligence, Artificial Intelligence, Machine Learning, Clinical Decision Support System, Human-Algorithm Collaboration, Multiple Sclerosis.

# REMERCIEMENTS

À mon directeur de thèse Prof. Franck MARS : Merci pour la confiance que vous m'avez témoignée et vos retours très réactifs et pertinents. Je vous remercie également pour le temps considérable que vous avez consacré à la relecture de mes articles. Vos encouragements et conseils m'ont été très précieux.

À mon co-directeur de thèse Prof. Pierre-Antoine GOURRAUD : Merci pour la confiance que vous avez accordée à ma modeste personne pour mener cette thèse. Merci d'avoir continuellement souligné la qualité de mon travail, cela m'a beaucoup encouragé à faire de mieux en mieux. Je vous remercie également pour l'enthousiasme débordant que vous avez manifesté pour ce projet et pour votre générosité scientifique, qui m'ont énormément aidé dans la réalisation de ce travail.

À mon encadrant de thèse Prof. Cédric DUMAS : Merci pour votre encadrement, vos conseils et vos nombreux points de vue constructifs, qui m'ont été indispensables pour progresser durant ces trois années. Merci pour votre présence et surtout votre patience. Au-delà de l'encadrement de cette thèse, vous avez été pour moi un bon grand frère.

Aux rapporteuses et aux membres de mon jury : Merci à Madame Céline LOUAPRE et Madame Isabelle PECCI d'avoir accepté d'être les rapporteuses de ce travail. Je tiens à vous exprimer toute ma gratitude. Merci à Monsieur Gilles EDAN et à Monsieur Pascal SEMPÉ d'avoir accepté de faire partie du jury. C'est un grand plaisir de vous avoir tous en tant qu'évaluateurs de ce travail, qui m'a passionné, et qui j'espère sera à la hauteur de vos attentes.

Au Prof. Gilles EDAN : Vous êtes le grand-père spirituel du travail effectué dans ce projet de thèse. Je tiens à vous remercier pour le partage constant de vos connaissances, pour vos conseils très constructifs et votre disponibilité. La confiance que vous m'avez témoignée et votre enthousiasme ont été déterminants dans la réalisation de cette thèse. Pour moi, vous êtes un excellent exemple de médecin qui cherche avec toutes les armes le bien-être des patients.

Au Prof. Gaël VAROQUAUX : Je tiens à vous remercier pour le temps que vous avez consacré au suivi de ma thèse durant ces trois années. Les retours, orientations et conseils que vous m'avez fournis me sont très précieux.



Aux personnels médicaux des centres hospitaliers de Rennes et de Nantes : Dr. Anne KERBRAT, Dr. David-Axel LAPLAUD, Dr. Emmanuelle LE PAGE, Prof. Emmanuelle LERAY, Dr. Michel LAURE, Mme. Nathalie BLANC, la liste est très longue, infirmier(e)s, neurologues, chercheur(e)s, chef(fe)s de projets, etc. Je tiens à vous remercier pour l'accueil, la collaboration, vos points de vue, votre aide et votre bienveillance. Sans vous, la réalisation de ce travail aurait été compliquée.

Aux patients : Il est vrai que je ne vous ai pas tous rencontrés, je connais la majorité dans les bases de données. Pourtant, vous êtes le pilier de ce travail, sans vous rien ne serait possible. Merci de m'avoir donné l'opportunité d'assister à vos consultations. Merci d'avoir accepté de participer à mon étude et aux études de collecte de données. Merci pour tout ce que j'ai appris grâce à vous. Merci également à Biogen et à l'OFSEP qui, par leur volonté de coopérer et de soutenir la recherche, m'ont permis d'accéder à vos données.

Aux personnels du LS2N et du CR2TI : Merci Annie et Haritiana pour votre gentillesse et votre aide au tout début de ma thèse, merci de m'avoir guidée et d'avoir répondu à toutes mes questions avec bienveillance, merci aussi pour votre beau sourire. Je tiens à remercier Virginie et Lisa qui ont pris la bonne suite de votre travail, nous n'avons pas eu la chance d'échanger de près, mais votre belle personne est visible dans vos mails. Merci à Angéline, Céline, Lucie et Sylvie d'être aussi sympathiques.

À tous mes collègues du CR2TI : Merci pour les échanges, les pauses, pour votre écoute et vos encouragements. J'en oublie, mais merci à Anaïs, Axelle, Éléa, Estelle, Florent, Irène, Joëlle, Léo, Magali, Olivia, Sonia, Stanislas et Vincent. J'espère vous retrouver pour de nouvelles expériences et je vous souhaite une très bonne continuation. Merci à Roxhaya et Sirine pour votre soutien, vos conseils et pour nos discussions, l'avenir sera meilleur. Un merci particulier à Ven, merci pour ta personne, tu étais là pour tout le monde. Merci pour ton soutien et ta bonne humeur. C'est toujours un plaisir de te retrouver et de voir que tu avances bien sur ton chemin.

À tous mes collègues de l'IMT-Atlantique et de l'équipe PACCE : Je tiens à vous remercier pour votre accueil chaleureux, votre soutien, vos conseils, votre partage et pour la sympathie dont vous avez fait preuve. Merci pour les moments que nous avons passés ensemble et j'espère que ce ne sont pas les derniers. J'en oublie forcément, mais merci à Agathe, Agus, AurélienL, Charlotte, Elise, Fida, Marine, Mathis, Mathieu et Yasmine

pour nos échanges et discussions sur tout et rien autour d'un café et des délices préparés par notre chef AurélienM ou autour d'un verre au COCO. Alex, bien que tu sois parti, il est vraiment appréciable que tu continues à rester en contact avec nous et à nous offrir ton aide. Merci pour ta belle humanité. Merci Celia pour ta belle personne et pour les merveilleux moments que nous avons passés ensemble au bureau et aussi avec tes formidables copains, que je tiens à remercier profondément. Merci Arnaud d'être là de temps en temps pour nous déranger avec tes tests, mais tu restes une personne adorable. Liv, merci de partager tes histoires avec tes mignons enfants et d'être une source d'inspiration avec tes motivations. Merci coach AurélienM pour ton écoute. Aylen, tu as rejoint l'équipe récemment, mais il me semblait que tu étais là depuis longtemps, tu es une très belle rencontre. Je salue au passage le talentueux et impressionnant Laurent et tous les autres basketteurs du mercredi et du dimanche, les matchs avec vous vont beaucoup me manquer.

À ma famille et mes amis : Merci maman et papa, les mots me manquent pour exprimer toute ma gratitude pour tout ce que vous avez fait pour moi, merci d'être mes parents. Merci à ma sœur et mon frère pour votre écoute, vos encouragements et votre soutien à toutes mes décisions. Merci à toute la famille EIOUAI et KHAL pour votre grand soutien et pour prendre des nouvelles régulièrement malgré les distances. Merci à toutes les autres personnes qui me sont proches : ma grand-mère, mes tantes, cousin(e)s et ami(e)s. Merci à Fatiha et Mina pour vos conseils, votre accueil et votre écoute. Merci aux petits Abd Arrahman, Anas, Basma, Lilia, Rimas, Ritaj et Yanis pour votre sourire et votre amour qui me donnent la joie de vivre et le courage d'avancer. Merci Hasnaa pour ta présence depuis tout ce temps, pour m'avoir écouté, merci d'être ma confidente. Sans oublier Geo, Maroua, et Ramiz, merci pour les bons moments passés avec vous qui j'espère ne seront pas les derniers.

À Ali : Merci d'être toujours présent, pour me soutenir, m'accompagner et m'encourager. Merci de me demander de n'être rien d'autre que moi-même. Merci d'avoir croisé mon chemin et d'avoir fait le pari fou de partager mon univers.

Enfin, je remercie tous celles et ceux qui ont contribué, directement ou indirectement, à faire de cette thèse une véritable aventure scientifique et humaine.



# ACRONYMES

|                |   |
|----------------|---|
| <b>ANR</b>     | Agence Nationale de la Recherche.                         |
| <b>AUC</b>     | Area Under the ROC Curve.                                 |
| <b>BAGGING</b> | Bootstrap AGGREGatING.                                    |
| <b>BOC</b>     | Bandes Oligoclonales.                                     |
| <b>CDSS</b>    | Clinical Decision Support System.                         |
| <b>CHU</b>     | Centre Hospitalier Universitaire.                         |
| <b>DIS</b>     | Dissemination In Space.                                   |
| <b>DIT</b>     | Dissemination In Time.                                    |
| <b>EBV</b>     | Epstein-Barr Virus.                                       |
| <b>EDSS</b>    | Extended Disability Scale Score.                          |
| <b>FDA</b>     | Food and Drug Administration.                             |
| <b>FLAIR</b>   | Fluid Attenuated Inversion Recovery.                      |
| <b>FN</b>      | Faux Négatif.   |
| <b>FP</b>      | Faux Positif.   |
| <b>GAMLSS</b>  | Generalized Additive Model for Location, Scale and Shape. |
| <b>HIPAA</b>   | Health Insurance Portability And Accountability Act.      |
| <b>IA</b>      | Intelligence Artificielle.                                |
| <b>IHM</b>     | Interaction Humain-Machine.                               |
| <b>IRB</b>     | Institutional Review Board.                               |
| <b>IRM</b>     | Imagerie par Résonance Magnétique.                        |
| <b>LCR</b>     | Liquide Céphalo-Rachidien.                                |
| <b>LIME</b>    | Local Interpretable Model-agnostic Explanations.          |
| <b>LMS</b>     | Lambda-Mu-Sigma.  |
| <b>ML</b>      | Machine Learning.   |

|                |  |
|----------------|--|
| <b>MP</b>      | Médecine de Précision.                           |
| <b>MSIF</b>    | Multiple Sclerosis International Federation.     |
| <b>OFSEP</b>   | Observatoire Français de la Sclérose En Plaques. |
| <b>POI</b>     | Patient Of Interest.                             |
| <b>POR</b>     | Patient Of Reference.                            |
| <b>RCT</b>     | Randomized Controlled Trial.                     |
| <b>RGPD</b>    | Règlement Général sur la Protection des Données. |
| <b>ROC</b>     | Receiver Operating Characteristic.               |
| <b>SCI</b>     | Syndrome Cliniquement Isolé.                     |
| <b>SEP</b>     | Sclérose En Plaques.                             |
| <b>SEP-RR</b>  | Sclérose En Plaques Récurrente-Rémittente.       |
| <b>SEP-PS</b>  | Sclérose En Plaques Progressive Secondaire.      |
| <b>SEP-PP</b>  | Sclérose En Plaques Primaire Progressive.        |
| <b>SF</b>      | Système Fonctionnel.                             |
| <b>SHAP</b>    | SHapley Additive exPlanations.                   |
| <b>SNC</b>     | Système Nerveux Central.                         |
| <b>SRI</b>     | Syndrome Radiologiquement Isolé.                 |
| <b>SVC</b>     | Support Vector Classifier.                       |
| <b>VN</b>      | Vrai Négatif.                                    |
| <b>VP</b>      | Vrai Positif.                                    |
| <b>XGBoost</b> | Extreme Gradient Boosting.                       |

# PRODUCTIONS SCIENTIFIQUES

## Article dans une revue :

1. Chadia Ed-Driouch, Franck Mars, Pierre-Antoine Gourraud, Cédric Dumas. Addressing the Challenges and Barriers to the Integration of Machine Learning into Clinical Practice : An Innovative Method to Hybrid Human–Machine Intelligence. *Sensors*, 2022, 22 (21), pp.8313. [1]
2. Chadia Ed-driouch, Florent Chéneau, Françoise Simon, Guillaume Pasquier, Benoit Combès, et al. Multiple sclerosis clinical decision support system based on projection to reference datasets. *Annals of Clinical and Translational Neurology*, 2022, 9 (12), pp.1863-1873. [2]

## Article de conférence :

1. Chadia Ed-Driouch, Pierre-Antoine Gourraud, Cédric Dumas, Franck Mars. The Integration of Human Intelligence into Artificial Intelligence to Provide Medical Practice-Based Predictions. *HHAI2022 : Augmenting Human Intellect*, Juin 2022, Amsterdam, Pays-Bas, vol. 354, pp. 293–294. [3]
2. Chadia Ed-Driouch, Gilles Edan, Franck Mars, Cédric Dumas, Pierre-Antoine Gourraud. L'interaction Homme-Machine au Service de la Médecine de Précision. Colloque du programme Data Santé, Machine quelle est ton humanité?, Juin 2022, Nantes, France. [4]
3. Chadia Ed-Driouch, Franck Mars, Pierre-Antoine Gourraud, Cédric Dumas. L'intelligence hybride pour prédire l'évolution des maladies chroniques. *IHM 2022 : 33ème conférence internationale Francophone sur l'Interaction Humain-Machine*, Avril 2022, Namur, Belgique. [5]

## **Poster de conférence :**

1. Chadia Ed-Driouch, Pierre-Antoine Gourraud, Cédric Dumas, Franck Mars. The Integration of Human Intelligence into Artificial Intelligence to Provide Medical Practice-Based Predictions. The first International Conference on Hybrid Human-Artificial Intelligence, Juin 2022, Amsterdam, Pays-Bas.
2. Chadia Ed-Driouch, Cédric Dumas, Franck Mars, Pierre-Antoine Gourraud. De la prise de décision à la communication avec le patient : mode dual de visualisation pour la gestion de la sclérose en plaque. 5ème journée Interaction Humain-Machine et Intelligence Artificielle, Mars 2021, Conférence en ligne, France.

## **Communication dans une journée scientifique :**

1. Pierre-Antoine Gourraud et Chadia Ed-Driouch. MS Vista, système d'aide à la décision clinique pour la sclérose en plaques : Projection de données individuelles dans un contexte personnalisé de références. Workshop, Systèmes d'aide à la décision-médecine-IA, Juin 2022, Nantes, France.
2. Pierre-Antoine Gourraud et Chadia Ed-Driouch. EUniWell, How Does Your Research Relate to the Topic of Well-Being?. Video Interview, France.

## **Poster dans une journée scientifique :**

1. Chadia Ed-Driouch, Franck Mars, Pierre-Antoine Gourraud, Cédric Dumas. Interaction humain machine pour l'aide à la décision médicale. Journée, L'I.A.nantaise s'invite au muséum!, Avril 2022, Nantes, France.

# TABLE DES MATIÈRES

|   |              |
|---|--------------|
| <b>Résumé</b>   | <b>v</b>     |
| <b>Abstract</b>   | <b>ix</b>    |
| <b>Remerciements</b>  | <b>xi</b>    |
| <b>Acronymes</b>  | <b>xv</b>    |
| <b>Productions scientifiques</b>  | <b>xvii</b>  |
| <b>Table des figures</b>  | <b>xxiii</b> |
| <b>Liste des tableaux</b>   | <b>xxv</b>   |
| <b>I Introduction</b>   | <b>1</b>     |
| I.1 Préambule . . . . .   | 2            |
| I.2 Contexte de travail . . . . .   | 2            |
| I.3 Problématiques scientifiques . . . . .  | 3            |
| I.4 Contributions visées . . . . .  | 5            |
| I.5 Structure du manuscrit . . . . .  | 6            |
| <b>II État de l’art général</b>   | <b>9</b>     |
| II.1 De la donnée clinique à la médecine de précision . . . . .   | 10           |
| II.1.1 Données cliniques pour l’aide à la décision . . . . .  | 10           |
| II.1.2 Médecine de précision, médecine personnalisée, médecine strati-<br>fiée, ou bien médecine factuelle? . . . . . | 13           |
| II.1.3 Applications, obstacles et défis de la médecine de précision . . . . .   | 15           |
| II.2 Du dialogue à la décision : principes et support pour la mise en place de la<br>décision clinique . . . . .      | 17           |
| II.2.1 Communication médecin-patient . . . . .  | 17           |
| II.2.2 Système d’aide à la décision clinique . . . . .  | 22           |
| II.3 Intelligence artificielle en médecine . . . . .  | 26           |



|            |   |           |
|------------|---|-----------|
| II.3.1     | Apprentissage automatique en médecine . . . . .   | 26        |
| II.3.2     | Adoption de l'apprentissage automatique dans les services de santé                          | 27        |
| <b>III</b> | <b>Sclérose en plaques</b>  | <b>31</b> |
| III.1      | Qu'est-ce que c'est la sclérose en plaques? . . . . .                                       | 32        |
| III.2      | Prévalence . . . . .  | 32        |
| III.3      | Origines et facteurs de risques . . . . .   | 34        |
| III.4      | Formes . . . . .  | 35        |
| III.5      | Symptômes . . . . .   | 37        |
| III.6      | Diagnostic . . . . .  | 38        |
| III.7      | Pronostic . . . . .   | 42        |
| III.8      | Traitements et suivi . . . . .  | 44        |
| III.9      | Enjeux et responsabilités vis-à-vis de la SEP . . . . .                                     | 48        |
| <b>IV</b>  | <b>Projection de données pour l'aide à la décision dans la SEP</b>                          | <b>51</b> |
|            | Introduction . . . . .  | 52        |
| IV.1       | État de l'art . . . . .   | 52        |
| IV.2       | Hypothèse . . . . .   | 53        |
| IV.3       | Enjeux des données . . . . .  | 53        |
| IV.4       | Outil de médecine de précision pour la SEP : MS Vista . . . . .                             | 54        |
| IV.4.1     | Analyse de l'activité . . . . .   | 55        |
| IV.4.2     | Conception . . . . .  | 57        |
| IV.4.2.1   | Contextualisation : POI / POR . . . . .   | 58        |
| IV.4.2.2   | Courbe d'évolution . . . . .  | 58        |
| IV.4.2.3   | Projection . . . . .  | 60        |
| IV.4.3     | Prototypage . . . . .   | 62        |
| IV.4.3.1   | Évaluation en vue d'un développement basé sur l'utilisateur                                 | 62        |
| IV.4.3.2   | Fonctionnalités de MS Vista . . . . .   | 63        |
| IV.4.4     | Évaluation de MS Vista comme outil support de la communication<br>médecin patient . . . . . | 71        |
|            | Discussion . . . . .  | 74        |

|  |                |
|--|----------------|
| <b>V Prédiction hybride pour l'aide à la décision médicale</b>                                 | <b>77</b>      |
| Introduction . . . . .   | 78             |
| V.1 État de l'art . . . . .  | 78             |
| V.2 Hypothèse . . . . .  | 79             |
| V.3 Nouvelle approche de l'intelligence hybride . . . . .                                      | 80             |
| V.3.1 Collaboration médecin-algorithme . . . . .   | 80             |
| V.3.2 Évaluation de l'approche d'intelligence hybride médecin-algorithme                       | 81             |
| V.3.3 Classes déséquilibrées : pipeline pour la création de modèles de<br>prédiction . . . . . | 82             |
| V.3.3.1 Séparation des données . . . . .   | 83             |
| V.3.3.2 Prétraitement des données . . . . .  | 83             |
| V.3.3.3 Classifieurs et réglage des hyperparamètres . . . . .                                  | 85             |
| V.3.3.4 Évaluation des performances . . . . .  | 89             |
| V.3.3.5 Méta-évaluation . . . . .  | 90             |
| V.3.3.6 Meilleur classifieur . . . . .   | 90             |
| V.4 Résultats . . . . .  | 90             |
| V.4.1 Modèles de prédiction de l'aggravation de l'EDSS . . . . .                               | 90             |
| V.4.2 Modèles de prédiction de l'aggravation des lésions T2 . . . . .                          | 91             |
| Discussion . . . . .   | 92             |
| <br><b>Conclusion générale et perspectives</b>   | <br><b>95</b>  |
| <br><b>Références</b>  | <br><b>101</b> |
| <br><b>Annexes</b>   | <br><b>125</b> |
| <br><b>Annexe A : Figures et tableaux annexes</b>  | <br><b>126</b> |
| <br><b>Annexe B : Documents utilisés dans les évaluations de MS vista</b>                      | <br><b>160</b> |



# TABLE DES FIGURES

|       |   |    |
|-------|---|----|
| II.1  | Nombre de publications par an utilisant les termes « médecine personnalisée », « médecine de précision » et « médecine stratifiée », publiées dans PubMed de 1997 à janvier 2019 (source J-T. Jørgensen [31]) . . . . . | 15 |
| III.1 | Cartes montrant la variation géographique de la prévalence de la SEP et des scores de confiance des données par pays (source : C. Walton et al. [105])  | 33 |
| III.2 | Quatre localisations préférentielles pour les lésions inflammatoires (source : ARSEP – Le point sur la SEP [121]) . . . . .   | 39 |
| III.3 | Échelle EDSS-simplifiée (source : sep.g-station.com) . . . . .  | 43 |
| IV.1  | Exemple de contextualisation d'un patient atteint de SEP . . . . .  | 59 |
| IV.2  | Processus de projection de données . . . . .  | 61 |
| IV.3  | MS Vista, interface de contextualisation et de projection des données . . . . .   | 63 |
| IV.4  | Visualisation de l'évolution de l'EDSS d'un POI par rapport à ses PORs (au fil du temps) dans MS Vista (source Ed-Driouch et al. [2]) . . . . .   | 64 |
| IV.5  | Sélection automatique des PORs dans MS Vista par des filtres prédéfinis de catégories de sélection des PORs . . . . .   | 65 |
| IV.6  | Sélection personnalisée des PORs dans MS Vista permettant de calibrer les catégories de sélection des PORs . . . . .  | 66 |
| IV.7  | Statistiques descriptives fournies par MS Vista sur les critères de sélection des PORs . . . . .  | 67 |
| IV.8  | Prévision de l'évolution potentielle de la SEP via les projections de données dans MS Vista . . . . .   | 68 |
| IV.9  | Mode dual de visualisation de l'aggravation de l'EDSS dans MS Vista . . . . .   | 69 |
| IV.10 | Illustration comparative des effets potentiels de chaque option thérapeutique dans MS Vista (source Ed-Driouch et al. [2]) . . . . .  | 70 |
| V.1   | L'intégration du raisonnement des médecins dans le processus de création du modèle de prédiction (source Ed-Driouch et al.[1]) . . . . .  | 81 |

|     |  |     |
|-----|--|-----|
| V.2 | Pipeline pour la création et la sélection des modèles de prédiction (source Ed-Driouch et al.[1]) . . . . .  | 84  |
| V.3 | Modèles prédictifs sélectionnés pour chaque type de modèle de prédiction de l'EDSS avec leurs scores AUC et F1 (source Ed-Driouch et al.[1]). . . . .      | 91  |
| V.4 | Modèles prédictifs sélectionnés pour chaque type de modèle de prédiction des lésions T2 avec leurs scores AUC et F1 (source Ed-Driouch et al.[1]). . . . . | 92  |
| A.1 | Graphe par paire montrant la façon dont les facteurs utilisés par le médecin diffèrent en fonction de l'EDSS . . . . .                                     | 126 |
| A.2 | Graphe par paire montrant la façon dont les facteurs utilisés et catégorisés par le médecin diffèrent en fonction de l'EDSS . . . . .                      | 127 |
| A.3 | Graphe par paire montrant la façon dont les facteurs utilisés par le médecin diffèrent en fonction des nouvelles lésions . . . . .                         | 128 |
| A.4 | Graphe par paire montrant la façon dont les facteurs utilisés et catégorisés par le médecin diffèrent en fonction des nouvelles lésions. . . . .           | 129 |

# LISTE DES TABLEAUX

|   |     |
|---|-----|
| III.1 Critères diagnostiques de McDonald 2017. . . . .  | 41  |
| III.2 Traitements de fond de la SEP (source : ARSEP – Le point sur la SEP [121])  | 45  |
| IV.1 Principaux indicateurs du contrôle thérapeutique de la SEP . . . . .   | 55  |
| IV.2 Catégories des facteurs de la SEP pour la contextualisation des patients . . .   | 56  |
| IV.3 Précisions sur la durée de l'évaluation . . . . .  | 72  |
| V.1 Proportion de chaque classe pour les facteurs à prédire . . . . .   | 83  |
| V.2 Explication de l'aggravation des indicateurs à prédire . . . . .  | 83  |
| V.3 Description des scores utilisés pour mesurer les performances des modèles<br>de prédiction. . . . .   | 88  |
| A.1 Statistiques descriptives des données quantitatives . . . . .   | 130 |
| A.2 Description des algorithmes de ML utilisés et du rôle de certains de leurs<br>paramètres clés [178–180] . . . . .                                       | 132 |
| A.3 Performance des modèles créés par la méthode <i>All Features</i> pour la prédic-<br>tion de l'aggravation de l'EDSS . . . . .                           | 141 |
| A.4 Performance des modèles créés par la méthode <i>Physician Features</i> pour la<br>prédiction de l'aggravation de l'EDSS . . . . .                       | 144 |
| A.5 Performance des modèles créés par la méthode <i>Physician Features &amp; catego-<br/>ries</i> pour la prédiction de l'aggravation de l'EDSS . . . . .   | 147 |
| A.6 Performance des modèles créés par la méthode <i>All Features</i> pour la prédic-<br>tion de l'aggravation des lésions . . . . .                         | 150 |
| A.7 Performance des modèles créés par la méthode <i>Physician Features</i> pour la<br>prédiction de l'aggravation des lésions . . . . .                     | 153 |
| A.8 Performance des modèles créés par la méthode <i>Physician Features &amp; catego-<br/>ries</i> pour la prédiction de l'aggravation des lésions . . . . . | 156 |



# CHAPITRE I

---

## INTRODUCTION

### Sommaire du chapitre

|     |  |   |
|-----|--|---|
| I.1 | Préambule . . . . .                    | 2 |
| I.2 | Contexte de travail . . . . .          | 2 |
| I.3 | Problématiques scientifiques . . . . . | 3 |
| I.4 | Contributions visées . . . . .         | 5 |
| I.5 | Structure du manuscrit . . . . .       | 6 |



## I.1 Préambule

La disponibilité de grandes quantités de données sur les patients et les maladies (cohortes, registres de maladies, essais cliniques, etc.) et les avancées scientifiques et technologiques (serveurs de calcul et de stockage puissants, intelligence artificielle, etc.) fournissent plus que jamais une meilleure opportunité d'avoir une vision plus précise des maladies [6]. Ces progrès constituent un contexte idéal pour développer des solutions de médecine de précision afin d'aider les professionnels de la santé à prendre des décisions cliniques individualisées pour chaque patient. Cependant, le manque de transparence et de compréhension de plusieurs de ces solutions constitue un obstacle majeur à leur adoption [7]. Il est important de s'assurer que ces solutions sont particulièrement claires en termes de quantité, de qualité et d'utilité clinique des données utilisées [8], mais aussi en termes de processus pour aboutir aux résultats suggérés [9].

Le dialogue humain-machine pour l'aide à la décision médicale tente de rendre la compréhension des données et de leur traitement la plus claire possible pour leur utilisateur [10]. Dans cette perspective, cette thèse propose des mécanismes de dialogue (visualisation de l'information, interactions avec les données ou les algorithmes) basés sur des concepts de la médecine de précision, afin de permettre une compréhension suffisante du fonctionnement et des bases des solutions développées pour soutenir le professionnel de santé et/ou le chercheur dans l'individualisation de la décision clinique. Le but est de maintenir leur confiance en établissant les bases d'une collaboration humain-machine fonctionnelle.

## I.2 Contexte de travail

Cette thèse est le fruit d'une collaboration entre Nantes Université, le CHU de Nantes et l'IMT Atlantique. Elle s'inscrit au programme Horizon 2020 de l'Union européenne pour la recherche et l'innovation dans le cadre du projet EU-TRAIN<sup>1</sup>. Le projet EU-TRAIN collecte et centralise les données cliniques des patients afin de les exploiter par l'Intelligence Artificielle pour aider à la décision médicale. Cette thèse a bénéficié d'un financement gouvernemental géré par l'Agence Nationale de la Recherche dans le cadre du

---

1. <https://eu-train-project.eu/>

programme « Investissements d’avenir » avec la référence KTD-Innov [ANR-17-RHUS-0010] et de la bourse Région Pays de la Loire numéro 2019\_11235.

J’ai réalisé mes travaux de recherche au sein de deux équipes : l’équipe PACCE (Perception, Action, Cognition pour la Conception et l’Ergonomie) du Laboratoire des Sciences du Numérique de Nantes (LS2N), et l’équipe 5 - Immunogénomique Translationnelle en Autoimmunité et Transplantation - ATIP Avenir, du Centre de Recherche en Transplantation et Immunologie Translationnelle (CR2TI). Au sein de l’équipe 5, mon travail était principalement centré sur les données cliniques, la médecine de précision, l’aide au traitement et au suivi des patients. Quant à mon travail au sein de l’équipe PACCE, il portait sur les thématiques de l’ingénierie des systèmes cognitifs, de la collaboration humain-machine, de la visualisation de l’information, et du traitement algorithmique des données.

### I.3 Problématiques scientifiques

Depuis sa première évocation en 1997 [11], le concept de médecine de précision vise à mieux caractériser et segmenter les maladies afin d’individualiser les décisions cliniques. Au regard des progrès de la science des données, de l’informatique et des biotechnologies, de nouvelles perspectives de la médecine de précision ont vu le jour afin d’améliorer les soins de santé et d’apporter une vision plus précise des maladies souvent complexes. Diverses solutions d’aide à la décision clinique ont été développées pour fournir une aide réelle au diagnostic, au pronostic, au choix thérapeutique ou encore pour améliorer la relation médecin-patient. En effet, l’objectif est d’accélérer l’identification des pathologies et d’aider à leur gestion, en tenant compte de la particularité de chaque patient. Cependant, l’acceptation et l’adoption de ces solutions dans les services de santé restent limitées par des problèmes liés aux données et aux méthodes utilisées pour développer les systèmes/modèles servant de base à ces solutions. Un aperçu des principales problématiques est présenté dans les paragraphes suivants.

**Les problématiques liées aux données** - L’avancement de la recherche en médecine dépend en tout premier lieu des données cliniques. Pour accéder à de telles données sensibles, il est obligatoire de veiller à ce que la vie privée des patients soit respectée et protégée [12] en se conformant aux normes de protection des données (par exemple, le

Règlement Général européen sur la Protection des Données - RGPD [13]). Cependant, les processus actuels pour se conformer à ces obligations sont généralement longs et peuvent entraver la mise à disposition en temps opportun des données. D'autre part, les préoccupations concernant la fiabilité, la qualité et l'utilité clinique des données utilisées constituent un obstacle important à l'acceptabilité de toute solution fondée sur les données [14]. La qualité des résultats obtenus est directement liée à la qualité des données. Lorsque les données ne reflètent qu'un groupe de la population, les résultats de leur utilisation ne seront appropriés que pour ce groupe et ne seront pas fiables pour les autres groupes [15]. Lorsque les données sont incomplètes ou de qualité moyenne, la qualité des résultats peut être insuffisante et nuire au patient à travers des recommandations approximatives par exemple.

**Les problématiques liées aux méthodes d'aide à la décision** - Les méthodes d'aide à la décision peuvent être basées sur une interaction directe avec les données ou via des algorithmes d'apprentissage automatique [5]. La complexité de l'interaction avec ces méthodes et l'interprétation des résultats qu'elles suggèrent [7], ainsi que la non-transparence des processus sous-jacents [9] et de la qualité/quantité des données utilisées [8], représentent un obstacle majeur à leur acceptabilité. Un autre problème critique concernant certaines de ces méthodes est leur élaboration sans intégrer la pratique clinique [5], ce qui peut entraîner une déconnexion de la réalité médicale ou un manque de confiance quant à leur fiabilité [3].

**Les problématiques liées aux systèmes d'aide à la décision clinique (CDSS)** - Un CDSS a le potentiel de soutenir la décision clinique et de la rendre partagée entre le médecin et le patient, ce qui s'est avéré important pour améliorer la satisfaction et l'adhésion au traitement [16]. Cependant, un CDSS ne permettant pas un contrôle direct et ne reflétant pas le raisonnement clinique peut avoir des conséquences inquiétantes [17]. Un CDSS peut perturber le déroulement du travail du médecin s'il est conçu sans tenir compte des connaissances et des suggestions des utilisateurs [18]. Lorsque le CDSS ne présente pas des informations faciles à comprendre et à interpréter par le médecin, l'interaction avec le patient peut être limitée par le temps passé avec le système [19]. Un CDSS destiné aux patients peut facilement entraîner des risques et des problèmes d'interprétation [20].

## I.4 Contributions visées

Dans le cadre de cette thèse, nous soulignons l'importance de la collaboration humain-machine pour soutenir la prise de décision clinique individualisée et rendre plus évidente la compréhension des données et leur traitement.

Notre première contribution porte sur l'utilisation de données médicales pour développer un système d'aide à la décision médicale dans la SEP. Le système devrait intégrer des principes de la médecine de précision et des méthodes basées sur le raisonnement clinique pour contextualiser les patients atteints de SEP. Il est censé permettre le calcul à la demande afin d'obtenir les informations nécessaires pour aider à prendre les bonnes décisions au bon moment. Le système devrait également être transparent en termes de données utilisées en permettant au clinicien d'examiner la qualité de la base de données dans son ensemble et la qualité des données utilisées pour contextualiser chaque patient et décider si celles-ci sont conformes et appropriées pour la prise de décision. Grâce à la projection des données, le système devrait illustrer les avantages potentiels de chaque option thérapeutique. Afin de permettre une compréhension suffisante du système et d'aider à la communication médecin-patient, il faudrait que le système soit doté de deux niveaux de visualisation. Le premier étant destiné au médecin pour expliquer les détails des données et des concepts utilisés. Le second pour enrichir la communication avec le patient et cibler les interactions autour des effets potentiels des différentes options thérapeutiques.

La deuxième contribution consiste à développer une nouvelle approche de prédiction hybride exploitant la puissance de calcul disponible, des données cliniques de bonne qualité, des algorithmes d'apprentissages automatique (ML - *Machine Learning*) et l'interaction avec des médecins experts. L'objectif est de concevoir des modèles de ML pertinents et compréhensibles par les humains afin de promouvoir l'adoption de systèmes basés sur le ML dans les services de santé. La validation attendue réside dans l'amélioration de la qualité des résultats de prédiction obtenus par l'approche hybride par rapport à l'approche conventionnelle de création de modèles de prédiction.

Les résultats des deux contributions précédentes ont été exploités dans un travail collaboratif avec d'autres membres de l'équipe 5 du CR2TI, pour développer une technique

collaborative d'anonymisation des données. Ma contribution a porté sur la préparation des données et leur catégorisation en fonction du raisonnement du médecin, afin que mes collègues puissent ensuite les utiliser pour alimenter les fonctions de création de données anonymes. Ce prétraitement des données est destiné à améliorer la qualité des résultats en évitant le risque de ré-identification des données tout en conservant leurs caractéristiques cliniques. Une autre contribution a été mise en place pour évaluer le système d'aide à la décision développé dans le cadre de la première contribution en tant que système d'aide à la communication médecin-patient. Cependant, les critères des patients qui peuvent participer à cette évaluation ont compliqué la réalisation de l'évaluation pendant la période autorisée pour l'utilisation des données.

### I.5 Structure du manuscrit

Ce manuscrit se compose de cinq chapitres. Le contenu de chaque chapitre est brièvement décrit dans les paragraphes suivants.

**Chapitre I : Introduction** - Ce chapitre présente un aperçu du contexte de travail, des problématiques scientifiques et des contributions envisagées.

**Chapitre II : État de l'art général** - Ce chapitre présente un état de l'art sur les thématiques liées à cette thèse. Il est composé de trois sections. La première est consacrée aux données cliniques et à la médecine de précision. Elle présente les données cliniques, la nécessité de les partager pour faire avancer la recherche translationnelle, et les enjeux liés à leur faisabilité et à leur qualité, ainsi qu'aux exigences légales et éthiques. Elle présente par ailleurs la différence entre la médecine de précision/personnalisée/stratifiée et la médecine factuelle, et se termine par les applications, les obstacles et les défis relatifs à la médecine de précision. La deuxième section présente un état de l'art sur la communication médecin-patient du point de vue de différentes perspectives. Elle décrit l'importance et les limites des systèmes d'aide à la décision médicale pour soutenir les décisions et les interactions en milieu clinique, et se conclut par les fondements de la conception d'un CDSS fiable. La troisième section représente les applications et les défis de l'apprentissage automatique en médecine et aboutit à des critères importants pour promouvoir l'adoption de systèmes d'apprentissage automatique dans les services de santé.

**Chapitre III : Sclérose en plaques** - Ce chapitre donne un aperçu général de la ma-

lady de la Sclérose En Plaques (SEP). Il présente la maladie, sa prévalence, son origine et ses facteurs de risque potentiels, ses formes et ses symptômes, les preuves cliniques permettant d'établir son diagnostic et son pronostic, ainsi que les options de traitement et de suivi. Il se termine par une discussion sur les enjeux et les responsabilités qui résultent de la SEP.

**Chapitre IV : Projection de données pour l'aide à la décision dans la SEP** - Ce chapitre présente la première contribution de cette thèse : MS Vista, un prototype fonctionnel de CDSS dans la SEP, fruit d'une collaboration avec des neurologues. Il représente un état de l'art des CDSS déjà développés pour aider à la prise de décision dans la SEP. Il décrit notre hypothèse pour l'aide à la décision dans la SEP. Il présente ensuite les données utilisées pour développer notre CDSS MS Vista et les différentes étapes de conception centrée sur l'utilisateur de son développement : analyse de l'activité, conception, prototypage et évaluation. Il décrit les concepts : contextualisation, POI, POR, projection des données, ainsi que les méthodes d'analyse et de visualisation des données, en particulier les percentiles. Il fournit des exemples de projection de données, de contextualisation et de visualisation. Les futures intégrations dans MS Vista dans le cadre du projet PRIMUS, pour lequel MS Vista a été utilisé comme preuve de concept pour obtenir son financement, sont présentées. Il se termine par une discussion sur la valeur ajoutée et les limites de MS Vista. Il se conclut par des exemples illustrant l'utilité de MS Vista dans l'individualisation des soins et des choix thérapeutiques.

**Chapitre V : Prédiction hybride pour l'aide à la décision médicale** - Ce chapitre présente la deuxième contribution de cette thèse : une nouvelle approche visant à combiner les intelligences humaine et artificielle pour soutenir la prise de décision médicale et promouvoir l'adoption de systèmes basés sur l'apprentissage automatique dans les services de santé. Il représente un état de l'art sur l'intelligence hybride et sa mise en œuvre pour créer des modèles de prédiction dans la SEP. Il décrit l'hypothèse initiale qui porte sur le bénéfice potentiel de la combinaison du raisonnement humain et artificiel. Il introduit les différentes étapes employées pour comparer la méthode conventionnelle de création de modèles de prédiction avec l'approche hybride proposée pour valider l'hypothèse en utilisant des données de patients atteints de SEP. Par ailleurs, il détaille les différentes étapes d'un pipeline de création de modèles prédictifs, développé dans le cadre de cette thèse pour remédier au problème de déséquilibre des classes qui est un problème typique

dans la SEP. Il présente ensuite une comparaison des performances des modèles de prédiction obtenus par chaque méthode. Enfin, il conclut par une discussion sur les résultats obtenus et le positionnement de la nouvelle approche d'intelligence hybride proposée par rapport à l'approche existante, et évoque les avantages de l'hybridation des intelligences humaine et artificielle.

Ce manuscrit se termine par une conclusion générale sur les travaux réalisés dans cette thèse et des perspectives sur des orientations et des évolutions possibles.

## CHAPITRE II

---

# ÉTAT DE L'ART GÉNÉRAL

### Sommaire du chapitre

|        |   |    |
|--------|---|----|
| II.1   | De la donnée clinique à la médecine de précision . . . . .  | 10 |
| II.1.1 | Données cliniques pour l'aide à la décision . . . . .   | 10 |
| II.1.2 | Médecine de précision, médecine personnalisée, médecine stratifiée, ou bien médecine factuelle? . . . . . | 13 |
| II.1.3 | Applications, obstacles et défis de la médecine de précision . . . . .                                    | 15 |
| II.2   | Du dialogue à la décision : principes et support pour la mise en place de la décision clinique . . . . .  | 17 |
| II.2.1 | Communication médecin-patient . . . . .   | 17 |
| II.2.2 | Système d'aide à la décision clinique . . . . .   | 22 |
| II.3   | Intelligence artificielle en médecine . . . . .   | 26 |
| II.3.1 | Apprentissage automatique en médecine . . . . .   | 26 |
| II.3.2 | Adoption de l'apprentissage automatique dans les services de santé . . . . .                              | 27 |



## II.1 De la donnée clinique à la médecine de précision

### II.1.1 Données cliniques pour l'aide à la décision

La recherche translationnelle favorise l'intégration multidirectionnelle de la recherche fondamentale, de la recherche orientée vers le patient et de la recherche sur les populations, dans le but à long terme d'améliorer la santé publique [21]. En effet, il s'agit de traduire les connaissances acquises en solutions pratiques et applicables pour les patients afin d'améliorer la prévention, le diagnostic, le traitement et la prise en charge des maladies. L'avancement de la recherche translationnelle dépend essentiellement de l'accès aux données des patients et aux connaissances cliniques. Le partage des données cliniques offre la perspective d'améliorer la reproductibilité et la pertinence de la recherche clinique. Ces données sont de plusieurs types [22], comprenant principalement :

- Les dossiers médicaux électroniques : les données recueillies dans un établissement médical (un hôpital, une clinique, etc.). Elles comprennent des données sur les patients, des données administratives et des données relatives au diagnostic, au traitement, aux médicaments prescrits, aux tests de laboratoire, à la surveillance physiologique, etc.
- Les registres de patients/de maladies : les registres de maladies sont des systèmes d'information clinique qui suivent un ensemble limité de données clés pour certaines maladies chroniques comme le cancer, le diabète et les maladies cardiaques. Ils fournissent des informations essentielles pour la gestion des maladies chroniques et de l'état des patients.
- Les enquêtes sur la santé : les enquêtes sur les maladies. Les plus courantes sont généralement menées pour fournir des estimations de prévalence afin d'obtenir une évaluation précise sur la santé de la population.
- Les données d'essais cliniques : un essai clinique est une étude scientifique menée sur des humains pour évaluer la non-toxicité, la sécurité et l'efficacité d'une méthode de diagnostic, d'un traitement ou d'une stratégie thérapeutique. Les essais cliniques se déroulent en plusieurs phases :
  1. L'étude de phase I : pour évaluer la sécurité et l'absence d'effets secondaires

chez des volontaires sains ou chez des patients en impasse thérapeutique (par exemple, des patients atteints de cancer).

2. L'étude de phase II : pour déterminer la dose optimale du médicament et ses éventuels effets secondaires.
3. L'étude de phase III : pour comparer le traitement à un placebo ou à un traitement de référence, afin de confirmer l'efficacité du nouveau traitement et de demander une autorisation de mise sur le marché.
4. L'étude de phase IV : vise à confirmer le bénéfice du nouveau traitement dans la population générale sur le long terme alors que le traitement est autorisé sur le marché.

La diffusion ouverte des données cliniques permettra aux chercheurs, cliniciens et autorités compétentes en matière de médicaments de les analyser et de les interpréter. Cela pourrait également permettre d'éviter la duplication des efforts dans la collecte des données, de réduire les coûts inutiles des études futures et d'encourager la collaboration au sein de la communauté scientifique [6]. La diffusion ouverte des données cliniques est toutefois confrontée à des défis opérationnels liés à des questions de faisabilité, de gouvernance, de durabilité et d'exigences juridiques/ éthiques différentes. L'accès aux données cliniques ne pouvant être direct, il est obligatoire de veiller à ce que la vie privée des patients soit respectée et protégée [12]. À cette fin, des réglementations et des mesures pour les utilisateurs ont été établies, notamment la loi américaine *Health Insurance Portability and Accountability Act* (HIPAA) de 1996 (loi qui sécurise et privatise les données relatives à la santé des citoyens américains) et le Règlement Général sur la Protection des Données (RGPD) de l'Union Européenne [23]. Les directives et les structures de gouvernance propres à chaque établissement, telles que les comités d'examen institutionnels (IRB - *Institutional Review Board*), sont également applicables à la recherche impliquant des patients et d'autres données sensibles (par exemple, les expérimentations sur des sujets humains, les données administratives, etc.). Les processus opérationnels actuels conçus pour répondre aux obligations peuvent entraver la livraison efficace et rapide des données, ce qui peut être particulièrement problématique dans le contexte des décisions où un accès prioritaire aux données est nécessaire pour éclairer la décision clinique. L'utilisation de données réelles est particulièrement prometteuse pour fournir une image exhaustive

de l'état du patient et aider à une meilleure prise de décision. Il convient au minimum d'obtenir les consentements appropriés et d'utiliser des techniques d'anonymisation des données pour garantir le respect des obligations en matière de confidentialité des données.

Les préoccupations concernant la fiabilité et la validité des données utilisées pour l'aide à la décision reste cependant un obstacle important à l'acceptabilité des solutions proposées [8]. La qualité des données est un aspect important des soins de santé et des technologies de l'information. Il s'agit de la mesure selon laquelle les données répondent aux attentes des utilisateurs et sont adaptées à leur objectif. Les technologies qui assistent les soins cliniques requièrent des données fiables et complètes, de sorte qu'une évaluation régulière de la qualité des données est nécessaire pour éviter les conséquences d'une mauvaise qualité. Les principales dimensions de la qualité des données qui ont fait l'objet de recherches dans le domaine médical [14] sont les suivantes :

1. L'exhaustivité : les données ne sont pas manquantes et sont d'une ampleur suffisante pour la raison recherchée.
2. L'exactitude : les données sont exactes et fiables.
3. La cohérence : les données sont conformes au type de modèle prévu.
4. L'accessibilité : la disponibilité et la facilité d'accès aux données.
5. L'actualité : les données sont suffisamment actuelles pour la tâche à accomplir.

Les études cliniques maximisent ces qualités, mais elles ne sont cependant pas entièrement garanties.

Les Essais Cliniques Randomisés (RCT - *Randomized Controlled Trial*) constituent actuellement la méthode de référence pour la certification des nouveaux moyens thérapeutiques [24]. Cependant, les critères d'inclusion et d'exclusion nécessairement stricts font que les populations étudiées ne sont pas forcément représentatives des populations de patients rencontrées dans la pratique clinique. Un partage judicieux des données d'essais cliniques pourrait aider à générer de nouvelles hypothèses de recherche sur la sécurité et l'efficacité des thérapies médicamenteuses. Leur utilisation avec des données en situation réelle pourrait aider à explorer de nouvelles questions et méthodes analytiques au-delà de celles incluses dans l'étude principale du RCT.

Les études en situation réelle sont un autre moyen important de comprendre l'utilité des interventions médicales dans une population de patients plus large et plus représentative [25]. En revanche, des informations importantes risquent de ne pas figurer systématiquement dans ces données. À titre d'exemple, lors de la réalisation de l'analyse d'activité avec un neurologue, j'ai assisté à une consultation où le médecin a demandé à la patiente (une jeune fille) si elle fumait, elle a confirmé et a demandé au médecin de ne pas l'écrire dans son compte rendu, comme elle ne voulait pas que ses parents le sachent. De la même manière, d'autres informations qui peuvent être importantes pour la décision médicale peuvent ne pas apparaître dans les données. Cela soulève la question de la fiabilité de l'utilisation des données issues des comptes rendus médicaux pour l'aide à la décision, et de la manière de surmonter ce problème.

Une approche unique de l'utilisation des données pour l'aide à la décision n'est pas suffisante. Par ailleurs, nous mettons en avant l'hypothèse que l'intelligence hybride représente une méthode prometteuse pour faire avancer la recherche translationnelle en proposant des solutions cohérentes et personnalisées, adaptées au contexte spécifique de chaque patient en fonction du problème de santé posé.

### **II.1.2 Médecine de précision, médecine personnalisée, médecine stratifiée, ou bien médecine factuelle ?**

La complexité de certaines maladies a nécessité le passage d'une approche centrée sur la maladie à une approche davantage centrée sur le patient, comme la Médecine de Précision (MP). Le terme « médecine de précision » a été évoqué pour la première fois en 1997 par P. Wasi [11] dans le cadre d'une discussion sur la révolution que la médecine génomique pourrait apporter au domaine de la biologie et de la santé en tant que base pour le développement de nouveaux outils de diagnostic, de pronostic, de traitement et de prévention. Deux ans plus tard, le terme « médecine personnalisée » a été introduit par R. Langreth et M. Waldholz [26] afin de souligner les avantages des médicaments sur mesure qui traitent les personnes en fonction de leur profil génétique individuel, par opposition à ce que les auteurs ont appelé *one-size-fits-all medicines* : les médicaments à composition unique pour tous les patients qui sont commercialisés par l'industrie pharmaceutique.

En 2005, le terme « médecine stratifiée » a été suggéré par S-X. Hu et ses coauteurs dans un article sur la médecine personnalisée et la pharmacogénomique (l'étude des effets du profil génétique d'un patient sur sa réponse aux médicaments) [27]. L'étude avait défini la médecine stratifiée comme application de la pharmacogénomique permettant de stratifier les patients en quatre segments :

1. Les répondeurs : patients qui répondent positivement à un médicament.
2. Les super-répondeurs : patients qui répondent particulièrement bien à un médicament. Plus efficacement que le segment (1).
3. Les non-répondeurs : patients qui ne répondent pas à un médicament en termes d'efficacité ou d'effets indésirables.
4. Les répondeurs négatifs : patients qui développent des effets indésirables inacceptables en réponse à un médicament.

La conclusion de l'étude fut que la médecine stratifiée était le terme le plus approprié à utiliser, car il reflète les effets réalistes des médicaments au niveau de la population, alors que le terme médecine personnalisée reflète plus précisément la promesse d'un ciblage et d'un développement uniques et individualisés des médicaments.

Depuis 1997, les termes médecine de précision, médecine personnalisée et médecine stratifiée ont été employés pour décrire l'individualisation des décisions médicales, contrairement à la médecine factuelle (*Evidence-Based Medicine*), qui cherche à déterminer le meilleur mode de prise en charge pour un patient en se basant sur des connaissances généralisables issues d'études de population (par exemple, l'effet moyen d'un traitement issu d'un essai randomisé) [28]. Le terme « médecine stratifiée » reflète plus ou moins la façon dont la pharmacothérapie individualisée est actuellement pratiquée, cependant, il est beaucoup moins utilisé. La médecine personnalisée ou la médecine de précision sont considérées par certains comme étant le même concept, sauf que le mot « personnalisée » a suscité des inquiétudes quant au fait que la médecine personnalisée puisse être interprétée à tort comme signifiant que des traitements entièrement personnalisés sont disponibles pour chaque patient unique, ce qui n'est pas le cas, d'où le passage au terme médecine de précision [29]. Le terme MP désigne l'adaptation du traitement aux facteurs individuels en regroupant des sous-populations de patients en fonction de leurs facteurs (cliniques, démographiques, génétiques, mode de vie, facteurs environnementaux, etc.)

afin de cibler les bons traitements au bon moment [30]. D'autres considèrent que la MP est bien une approche innovante visant à adapter la prévention et le traitement des maladies, et est également appelée médecine personnalisée. Comme l'illustre la Figure II.1, ces dernières années, le terme préféré est passé de médecine personnalisée à médecine de précision [31].

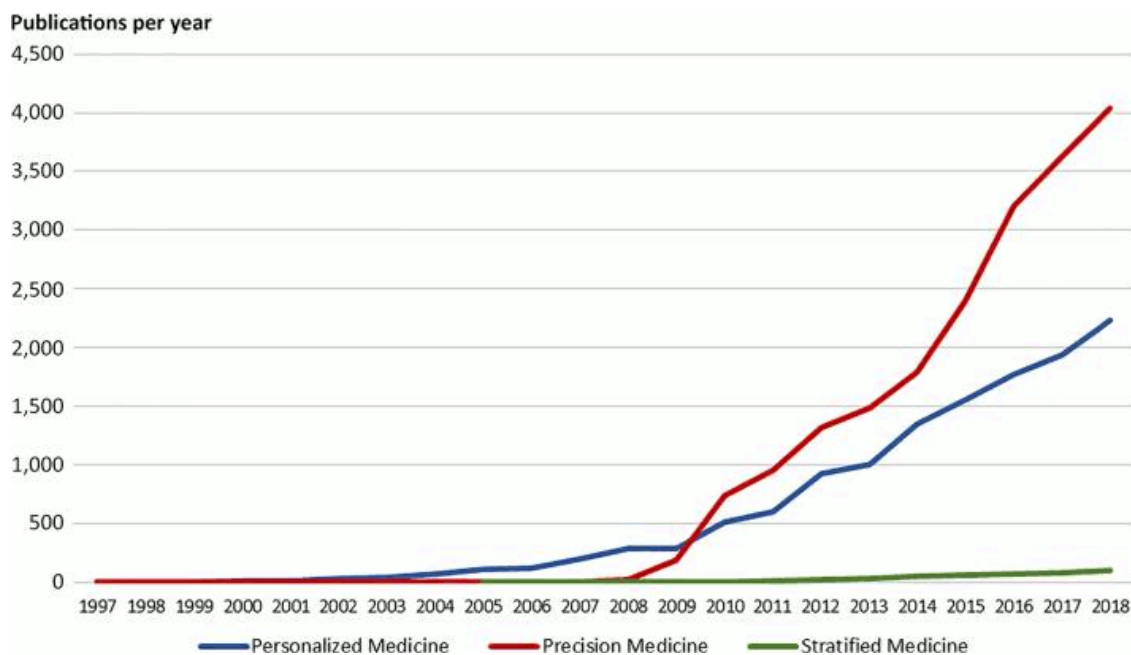


FIGURE II.1 – Nombre de publications par an utilisant les termes « médecine personnalisée », « médecine de précision » et « médecine stratifiée », publiées dans PubMed de 1997 à janvier 2019 (source J-T. Jørgensen [31])

Indépendamment du terme médecine de précision ou médecine personnalisée, les approches qui ont vu le jour il y a plus de deux décennies ont eu un impact considérable sur la manière dont les nouvelles stratégies de prise de décision, de pronostic, de diagnostic, de prévention et de traitement sont développées et mises en œuvre aujourd'hui.

### II.1.3 Applications, obstacles et défis de la médecine de précision

La MP est une approche fondée sur les données dont l'objectif est de personnaliser la prise décision clinique afin de délivrer des mesures spécifiques pour une prise en charge individualisée du patient permettant de fournir des traitements, diagnostics et mesures de prévention plus précises.

À la lumière des progrès de la biotechnologie avec la diminution des coûts de séquençage, la disponibilité de grandes quantités de données (cliniques, IRM, biologiques, génétiques, etc.) et de dispositifs permettant de fournir des flux de données à haute résolution en temps réel (ex, capteurs physiologiques du suivi télémedical) associés à des avancées en science des données et en informatique (les serveurs de stockage, les serveurs de calcul, l'informatique en nuage, l'internet des objets, la cybersécurité, etc.), de nouvelles voies de recherche ont émergé pour améliorer les soins de santé et pour une vision plus précise des maladies. Ces diverses avancées offrent un véritable support à la MP pour fournir des calculs à la demande et un accès sécurisé aux données, afin de proposer aux médecins différents protocoles ou projections pour un même patient en tenant compte de ses facteurs. La MP a prouvé son importance dans la communauté scientifique en identifiant des facteurs génomiques et en améliorant de nombreux aspects dans différents domaines médicaux tels que l'oncologie de précision, la gestion, la détection et le traitement du cancer [32–34], l'oncologie pédiatrique [35–37], la pharmacogénomique [38], et l'introduction de thérapies moléculaires ciblées révolutionnaires dans divers domaines médicaux [39].

Les objectifs à long terme de la MP sont de mieux caractériser et segmenter la maladie, de faire preuve d'une détection et d'un suivi des symptômes le plus tôt possible, d'identifier les personnes présymptomatiques, et de superviser l'évolution et la gestion de la maladie en la rendant moins coûteuse et avec moins d'effets indésirables [29]. L'amélioration de la MP nécessite de mieux identifier les raisons sous-jacentes de la variabilité de la réponse aux médicaments, le médicament ou l'association de médicaments susceptible d'être le(s) plus efficace(s) pour un patient donné, ainsi que de prendre en compte certains facteurs peu discutables tels que le passé, le milieu et le mode de vie du patient [40]. La pandémie de COVID-19 est l'exemple le plus récent et le plus éloquent de la nécessité pour la médecine de précision d'aller plus loin et plus vite.

La médecine de précision dépend d'une large participation de la population générale, notamment d'un grand nombre de patients, de professionnels de la santé, de chercheurs, etc. Pour assurer cette participation, la confiance, la confidentialité et des bénéfices pour les participants sont nécessaires. Les violations de données très médiatisées, la peur de la réidentification, les préoccupations d'ordre juridique, l'absence d'un système complet de technologies de la santé et le système d'enseignement médical dépourvu de modules destinés à enseigner l'intégration des solutions de la MP dans la pratique, constituent

des obstacles à la MP. Une bonne communication claire et honnête avec les participants et leur inclusion dans la gouvernance de la recherche sont essentielles pour établir la confiance. Les protections juridiques des données et les approches technologiques visant à garantir la sécurité des systèmes d'information telles que l'anonymisation et le chiffrement des données, le contrôle d'accès via la blockchain, etc. peuvent également améliorer la confiance et favoriser l'acceptabilité [41]. Par ailleurs, l'intégration de l'enseignement de la MP dans la formation médicale, en plus de la collaboration avec l'industrie pharmaceutique, peut aider à relever les défis de la MP et à accélérer ses progrès [29].

En conclusion, la médecine de précision, en conjonction avec l'évolution scientifique et la diffusion ouverte des données, pourrait éventuellement renforcer la base factuelle des décisions cliniques individualisées en tirant des conclusions solides et plus significatives sur les avantages et les risques d'une intervention clinique ou thérapeutique.

## **II.2 Du dialogue à la décision : principes et support pour la mise en place de la décision clinique**

L'utilisation des données en MP peut aider l'activité médicale à différents niveaux, que ce soit pour des chercheurs, des médecins experts, des médecins moins habitués à certaines pathologies, ou encore à la communication dans les équipes médicales ou avec les patients.

### **II.2.1 Communication médecin-patient**

La communication médecin-patient est une composante essentielle en pratique médicale, complexe par nature, et sa qualité est plus ou moins liée à la qualité et l'efficacité des soins de santé. Les principaux objectifs de l'interaction entre le médecin et le patient sont le partage d'informations pour soutenir le diagnostic et le traitement, ainsi que le renforcement de la participation active des patients à leurs soins [42]. Plusieurs études ont été réalisées pour étudier les différents aspects de la communication médecin-patient. Les sociologues observent que le succès de la communication médecin-patient dépend du fait que le médecin traite les patients, plutôt que seulement la maladie [43], car plu-



sieurs facteurs peuvent influencer cette communication. Par exemple, les patients issus des milieux dont le statut socio-économique est faible ont tendance à participer moins activement à leurs consultations en posant moins de questions et sont moins susceptibles de donner de l'information spontanément ou d'exprimer leurs opinions/émotions. Ces patients influent à la fois sur la façon dont ils communiquent avec le médecin, ainsi que sur la façon dont le médecin communique avec eux [44]. D'autres indicateurs peuvent influencer la qualité des consultations médicales comme la confiance en le médecin, le sentiment d'être respecté par le médecin, et la capacité du médecin à annoncer les mauvaises nouvelles et à écouter, etc. [45]. Il a été constaté qu'une communication ouverte est le facteur le plus important de la relation patient-médecin [46] et qu'une communication imprécise pouvait conduire à une sous-estimation de la gravité de la maladie parce que les patients ne signalent pas les problèmes importants liés au mode de vie [47]. D'autre part, des chercheurs indiquent que les médecins et les patients évaluent la communication lors des consultations médicales différemment et, en général, leurs évaluations ne concordent que légèrement [48]. Cela dit, la consultation ne se déroule pas toujours tel que le clinicien le pense. Par conséquent, les cliniciens doivent être conscients des différents facteurs qui peuvent influencer leurs relations et les processus décisionnels de leurs patients et tenter d'identifier et de traiter ces facteurs [49]. D'un point de vue philosophique, le développement de la médecine d'aujourd'hui est accouplé avec le développement des liens théoriques et concrets entre les vertus et les principes qui visent le bien du patient. La vertu (force morale avec laquelle l'être humain tend au bien) est mieux transmise par un modèle ou un mentor (conseiller expérimenté) [50]. Cobos et al. [51] ont réalisé une étude psychologique visant à fournir l'étiologie de l'hypertension de la blouse blanche et à décrire des méthodes permettant d'améliorer l'interaction médecin-patient. L'étude s'est fondée sur une revue de littérature des recherches ayant précédemment abordé ce sujet. Il en ressort que l'inclusion du patient dans le processus de la prise de décision clinique peut réduire l'anxiété du patient et améliorer l'adhésion au traitement.

Une consultation de bonne qualité peut améliorer les décisions, la qualité des visites, l'efficacité du traitement, l'efficience du service, la qualité des soins, la sécurité des patients et la satisfaction des médecins et des patients [52]. Les compétences de communication lors de l'interaction avec les patients doivent continuellement s'améliorer tout au long de la carrière professionnelle des médecins [53]. L'étude de Bahadori et al. [54] s'est

intéressée spécifiquement aux perspectives et aux expériences des patients et des médecins pour déterminer les différents facteurs qui peuvent affecter la qualité des consultations médicales. Ils ont réalisé une étude qualitative par entretiens (en face à face et par téléphone) avec 21 médecins et 27 patients. Les résultats ont conduit à trois dimensions principales pour mesurer la qualité des consultations médicales :

- La qualité de la structure
- La qualité des processus
- La qualité des résultats

La qualité de la structure est un concept de base pour améliorer la qualité des services médicaux. Elle décrit l'environnement de l'organisation des soins, comme les équipements ou les parcours, mais aussi la qualification du personnel ou les structures administratives. La qualité du processus est un concept à multiples facettes intégrant les notions de confiance, d'empathie, d'équité et de respect. La qualité perçue du résultat du service influence positivement la satisfaction du patient dans divers types de services hospitaliers. En conclusion, ces dimensions ont le potentiel d'améliorer la qualité de la consultation en renforçant les aptitudes et les techniques du milieu médical. En 1999, une vingtaine de responsables de l'enseignement médical et d'experts en communication ont élaboré le consensus de Kalamazoo pour définir les éléments essentiels de la communication médecin-patient. Ils se sont réunis au cours des années suivantes pour aboutir, en 2001, à la publication d'une liste de sept éléments pertinents qui constituent un cadre utile pour les programmes d'études et les normes axés sur la communication médecin-patient [55] :

1. Établir la relation médecin-patient : établir un partenariat avec le patient en prenant en considération l'individu dans son ensemble et pas seulement la maladie.
2. Ouvrir la discussion : établir une connexion personnelle.
3. Recueillir des informations : susciter les préoccupations ou les sentiments du patient.
4. Comprendre le point de vue du patient : demander le contexte spécifique au patient.
5. Partager des informations : donner des renseignements appropriés sur un diagnostic, un pronostic, traitement, etc.

6. Parvenir à un accord sur les problèmes et les procédures : vérifier la volonté du patient de suivre les procédures médicales.
7. Conclure : discuter de la prochaine visite.

Idéalement, la prise de décision partagée sous-tend chacune de ces composantes [55]. Menichetti et al. [56] ont également suggéré un ensemble de stratégies que le médecin peut choisir pour fournir des informations, mais qui peuvent être difficile à appliquer dans le monde réel de la pratique médicale, car leur application nécessite un temps important.

L'application des technologies de l'information et de la communication dans les services de santé pourrait également améliorer la qualité de la relation patient-médecin [54]. Hedian et al. [57] ont effectué une revue de la littérature sur l'adoption des dossiers électronique de santé et son effet sur les examens cliniques. Parmi les questions abordées, l'impact de l'introduction de l'ordinateur dans la relation médecin-patient. Il a été constaté qu'entre 1971 et 2016, le temps consacré aux interactions directes avec le patient a diminué de 14.5% à 9.4%. Cependant, l'ordinateur n'est pas forcément perçu par les patients comme une intrusion. Les patients/familles de patients ont signalé une augmentation significative de leur satisfaction générale de la manière dont les informations médicales leur étaient communiquées et de leur compréhension du diagnostic et du plan de traitement. En revanche, de nombreux cliniciens ont exprimé leur inquiétude quant au rôle de plus en plus dominant que joue l'ordinateur dans leurs interactions avec les patients. Certains voient que les médecins passent plus de temps à interagir avec l'ordinateur et à y saisir des données qu'à prodiguer des soins directs aux patients. Pour aider les professionnels de la santé à devenir plus productifs en utilisant des systèmes numériques, il est important de présenter les informations des patients d'une manière simple et plus signifiante, et s'assurer que les données restent privées, sécurisées et disponibles aux bonnes personnes et au bon moment [58].

L'ingénierie des systèmes cognitifs (discipline de conception de processus et de technologie pour l'intégration humain-système en utilisant des analyses d'activité basées sur la pratique) et la conception d'interface écologique sont des compléments importants aux approches de l'ergonomie pour soutenir la pratique fondée sur des preuves dans les soins de santé [10]. Les visualisations jouent un rôle très important dans la représentation des

systèmes cognitifs, elles guident les processus de réflexion et de la communication [59]. Le langage visuel est l'une des principales avenues utilisées pour communiquer des informations sur la santé à travers les barrières de l'alphabétisme et pour compléter ou remplacer les informations qui étaient auparavant présentées dans le texte. Les icônes, les graphiques, les vidéos et les animations sont des exemples courants du langage visuel utilisé pour simplifier la communication [60].

En guise de conclusion, la plupart des études ont montré l'importance de la participation des patients dans le processus de prise de décision médical [61] pour améliorer la satisfaction, l'observance du traitement [62], et aussi la qualité des soins fournis au patient. Beers et al. [16] ont défini 4 étapes essentielles pour la prise de décision participative (la prise de décision partagée entre le médecin et le patient) :

1. Informer les patients qu'une décision doit être prise.
2. Expliquer les options du traitement.
3. Identifier les valeurs et les objectifs du patient tout en soutenant le débat.
4. Prendre la décision et organiser le suivi.

Les interfaces humain-machine en tant qu'espace interactif [63] ont le potentiel de soutenir la prise de décision clinique et la rendre partagée entre le médecin et le patient. Les visuels interactifs qu'elles fournissent, offrent aux médecins et aux patients la possibilité de voir comment les différents facteurs contribuent à la santé. En outre, l'utilisation de tels visuels associés à des catégories de traitement peut enrichir la communication médecin-patient sur les avantages potentiels de diverses options de traitement [64].

En effet, la MP et les progrès scientifiques offrent de plus amples possibilités pour améliorer le bien-être des patients. Cependant, si le patient n'adhère pas aux procédures médicales suggérées, quel serait le fruit de tous ces efforts? Pour compléter le bénéfice, il serait important d'adopter des techniques de prise de décision participative impliquant le médecin et le patient. Les études évoquées précédemment ont montré l'importance de cette approche dans l'amélioration de la satisfaction et de l'adhésion au traitement. Pour y parvenir, il est possible de visualiser l'évolution potentielle de la maladie chez le patient selon les différentes options de traitement, en se basant sur l'évolution de celle d'autres patients présentant des facteurs de la maladie similaires, par exemple. Des visuels inter-

actifs à différents niveaux peuvent être utilisés. Un premier niveau peut par exemple être conçu pour que le médecin évalue et valide les résultats illustrés et un second niveau pour enrichir la communication médecin-patient sur les différentes options thérapeutiques afin de choisir ensemble le traitement le plus approprié.

## II.2.2 Système d'aide à la décision clinique

Les décisions cliniques sont souvent prises au contact du patient lors d'une consultation ou d'une visite de suivi. À ce moment critique de la décision, le clinicien est censé disposer de tous les facteurs médicaux et les connaissances liés à l'état du patient. Au cours de la consultation, il est mis au défi de gérer un historique des données du patient, de chercher la bonne façon de connaître les données personnelles qui peuvent influencer l'état du patient (mode de vie du patient, difficultés, etc.), de se référer aux connaissances, d'expliquer et de communiquer avec chaque patient en fonction de son état de santé, de son âge, de son niveau de compréhension, etc. Le médecin est avant tout un être humain, qui peut être induit en erreur en oubliant/négligeant certaines informations importantes, ce qui peut conduire à des décisions peu fiables. Il peut également trouver des difficultés à expliquer/échanger avec certains patients, ce qui peut influencer la relation médecin-patient et ainsi créer des conséquences involontaires telles que décrites dans la section précédente. Il est donc utile de disposer d'un système d'aide à la décision capable de prendre en compte toutes les données disponibles et de fournir des illustrations pour aider à la prise de décision et à faciliter la communication médecin-patient.

Un système d'aide à la décision clinique (CDSS – *Clinical Decision Support System*) est un aspect important des technologies de l'information en santé. L'objectif principal des CDSS est d'aider les cliniciens dans les processus de prise de décision. Reed T. Sutton et al. [18] avaient défini deux types de CDSS :

1. Le CDSS traditionnel comme une aide directe à la prise de décision clinique, dans lequel les facteurs individuels du patient sont adaptés à une base de connaissances cliniques informatisée afin de présenter des évaluations spécifiques au clinicien.
2. Les CDSS modernes comme un outil au point de service que les cliniciens utilisent pour combiner leurs connaissances avec les suggestions fournies par le CDSS.

Selon A. T. M. Wasylewicz et al. [65], pour comprendre et se familiariser avec les CDSS,

il est important de les catégoriser en se basant sur :

- La fonction du système
- Le modèle de conseil
- Le style de communication
- Le processus décisionnel sous-jacent
- L'Interaction Humain-Machine (IHM)

La fonction du système vise à déterminer « ce qui est vrai? » sur la base d'un ensemble de données saisies par l'utilisateur ou facilement disponibles, et « que faire? », par exemple pour aider au pronostic du patient. Le modèle de conseil est l'ensemble des approches actives ou passives utilisées pour donner des conseils. Le style de communication catégorise le CDSS en distinguant deux modèles :

1. Modèle de consultation : le système est un conseiller qui interroge et propose des solutions/ actions. Par exemple suggestion d'un traitement en fonction du diagnostic.
2. Modèle de critique : le système permet à l'utilisateur de décider par lui-même et l'alerte s'il y a un risque autour de la décision prise.

Le processus décisionnel sous-jacent représente les méthodes utilisées pour l'aide à la décision, que ce soit des modèles basés sur la pratique médicale, la littérature, des approches statistiques ou l'intelligence artificielle, etc. L'IHM désigne la communication et l'interaction entre l'utilisateur et la machine par le biais d'une interface utilisateur qui permet de transmettre des informations de la machine à l'humain de manière interactive en utilisant des transactions de type requête-réponse [66]. Les interfaces humain-machine [67] ont le potentiel de faciliter l'accès et la compréhension des approches de la médecine de précision. Elles permettent l'interaction avec les données en temps réel, afin d'obtenir les informations nécessaires pour prendre la bonne décision au bon moment.

Un CDSS est une combinaison de données et de connaissances conçue pour fournir aux professionnels de la santé et aux patients des informations portant sur des procédures médicales susceptibles d'aider à prendre une décision individualisée pour chaque patient. Les CDSS offrent un large éventail de fonctions. Ils peuvent intervenir comme support au diagnostic, au pronostic ou encore au choix thérapeutique pour accélérer l'identification de pathologies souvent complexes et aider à leur prise en charge. Le développement

des CDSS est de plus en plus marqué : analyse des données d'IRM (Imagerie par Résonance Magnétique) pour caractériser les tumeurs [68], aide à la décision médicale dans les maladies complexes [2], aide aux pharmaciens pour fournir à leurs clients des conseils factuels en toute simplicité [69], aide à la gestion des pandémies [70], etc. En tant qu'espace d'interaction et de visualisation des données, les CDSS ont le potentiel d'accroître l'adhésion aux procédures cliniques en enrichissant l'interaction médecin-patient. Ils sont également capables d'améliorer la sécurité des patients en agissant comme des systèmes d'alarme pour rappeler les événements médicaux liés aux médicaments ou à d'autres facteurs pathologiques. À titre d'exemple, le CDSS développé par S. Eslami et al. [71] a permis d'améliorer la sécurité des patients en réduisant le pourcentage d'hypoglycémies grâce à un protocole de surveillance de la glycémie en fonction des facteurs individuels de chaque patient.

Malgré les nombreux développements et efforts pour mettre en œuvre les CDSS, l'adoption et l'acceptation restent limitées par certaines difficultés telles que l'accès à des données de qualité approuvées pour le cas d'intérêt [14], les mesures de protection de la vie privée [12], la sécurité de ces systèmes, etc. Ainsi que des défis liés à l'impact des CDSS sur les professionnels qui les utilisent et leurs conséquences à l'égard des patients [18]. Les CDSS peuvent perturber les processus de travail des professionnels de la santé s'ils sont conçus sans tenir compte des connaissances humaines, s'ils n'intègrent pas la pratique médicale, ou s'ils ne permettent pas un dialogue entre le système et l'utilisateur, comme le cas des systèmes automatiques [72]. Et même si le CDSS est fondé sur les bases de connaissances nécessaires, lorsque l'utilisateur n'est pas suffisamment familiarisé avec le système ou avec les systèmes informatiques en général, les échanges avec le patient peuvent être limités par le temps que l'utilisateur passe à interagir avec le système [57]. Un autre problème critique des CDSS est la quantité de notifications ou d'alertes générées dont certaines sont sans importance ou parfois incorrectes que les cliniciens doivent souvent vérifier par eux-mêmes [19]. Ceci amène les utilisateurs à être moins attentifs et donc à négliger des alertes potentiellement très utiles. À long terme, les utilisateurs peuvent développer une forme de complaisance ou de surconfiance [73] à l'égard du CDSS, au point d'en perdre leur indépendance vis-à-vis du système et d'avoir des difficultés à effectuer certaines tâches s'ils passent à un environnement sans CDSS [18].

Un CDSS est le fruit d'une technologie évolutive qui a le potentiel d'aider à l'indi-

vidualisation de la décision médicale, en facilitant l'accès aux solutions de médecine de précision. Il peut également améliorer la relation médecin-patient en tant que support de communication. Pour promouvoir leur adoption, il est primordial de concevoir des CDSS intégrant les bases et les connaissances de la pratique médicale. Cela doit se faire sous la forme de règles d'inférence explicites qui soient adaptées aux capacités des utilisateurs et offrent des performances satisfaisantes, fiables et acceptables [74].

Musen et al. [75] et Osheroff et al. [76] ont présenté les règles de base pour construire un CDSS efficace comme suit :

- L'accès à des données de patients de bonne qualité, approuvées et valides pour leur usage clinique.
- L'accès à des connaissances pertinentes éclairant la pratique clinique.
- L'organisation et l'encapsulation des connaissances cliniques dans un modèle capable de soutenir correctement le processus de prise de décision.
- La conception de représentations facile à comprendre et à interpréter qui reflètent les connaissances et le raisonnement liés à l'usage clinique.

En conclusion, les CDSS offrent un ensemble de fonctionnalités pouvant servir à la prise de décision médicale individualisée. En combinant les données et les connaissances dans des visuels interactifs, les CDSS offrent une opportunité importante d'accéder et de comprendre les solutions de médecine de précision. Cela pourrait également ouvrir la voie à une prise de décision participative, renforçant la relation médecin-patient, comme indiqué à la fin de la section précédente. Cependant, des défis subsistent, tels que la qualité des données et des méthodes utilisées, les enjeux de confidentialité et de sécurité, et la surabondance d'alertes. Pour favoriser leur adoption, il est essentiel de concevoir des CDSS qui intègrent les connaissances issues de la pratique médicale, tout en offrant des performances fiables et acceptables. Il est également important que ces systèmes présentent les informations d'une manière visuellement claire et compréhensible, facilitant la communication entre le système et les professionnels de la santé, ainsi qu'entre ces derniers et les patients.



## II.3 Intelligence artificielle en médecine

Face aux grandes quantités de données disponibles sur les patients, notamment les cohortes et les essais cliniques, ainsi que le développement des infrastructures de stockage de données et les progrès de la science des données, l'Intelligence Artificielle (IA) est devenue un outil qui offre un environnement favorable pour relever le défi des décisions médicales individualisées [5].

### II.3.1 Apprentissage automatique en médecine

Les premières recherches sur l'utilisation de l'IA en médecine remontent aux années 1950 avec l'élaboration de règles de diagnostic et de traitement par des chercheurs comme Edward Shortliffe [77]. Le premier système d'IA, DENDRAL, a été développé pour interpréter les spectres de masse et identifier des molécules [78]. Depuis lors, l'IA médicale a connu une croissance exponentielle en termes de développement de techniques et d'applications. En 1972, MYCIN, un système de diagnostic et de traitement des infections bactériennes, a été développé par Edward Shortliffe et son équipe à l'Université de Stanford [79]. En 1974, le système Internist-I, a été développé à l'Université de Pittsburgh en tant qu'expérience éducative pour aider à diagnostiquer les maladies internes. Il utilisait des règles d'inférence pour identifier les maladies possibles en fonction des symptômes du patient [80]. Dans les années 1980 et 1990, des systèmes d'IA plus avancés ont été développés, notamment les réseaux de neurones artificiels qui ont commencé à être utilisés dans l'IA médicale [81]. En 1985, le système ONCOCIN a été développé pour aider les oncologues à établir des plans de traitement pour les patients atteints de cancer du poumon [82]. Autour de l'an 2000, la médecine personnalisée a commencé à émerger, en utilisant des algorithmes d'apprentissage automatique pour identifier des sous-groupes de patients qui sont plus susceptibles de répondre à certains traitements [83]. Plusieurs autres applications de l'IA en médecine ont vu le jour au fil des ans, en imagerie médicale, en génomique, en développement de médicaments, etc. [84]. Plus récemment, l'IA a été utilisée pour aider à la recherche et au développement de vaccins contre le COVID-19 [85].

On peut distinguer deux branches principales de l'IA en médecine, telles que définies

par P. Hamet et J. Tremblay [86] :

- Une branche physique : représentée par les objets physiques, les dispositifs médicaux et les robots utilisés pour assister le patient âgé ou le chirurgien.
- Une branche virtuelle : représentée par les algorithmes (par exemple, les algorithmes d'apprentissage automatique).

L'apprentissage automatique (ML - *Machine Learning*) est l'une des composantes essentielles des applications de l'IA dans le domaine des soins de santé. La flexibilité et la compétence des algorithmes d'apprentissage automatique, avec la puissance de calcul disponible, ont un rôle particulier à jouer dans l'avancement de la médecine personnalisée. Ils dépassent la capacité des individus à prédire en temps réel les futurs événements liés à la santé à partir des données médicales existantes [87]. Les algorithmes de ML permettent l'optimisation d'une fonction objective basée sur des données complexes et multidimensionnelles [7]. Ils sont capables de détecter des modèles, des différences et des corrélations dans les données pour faire des prédictions sur la probabilité de résultats incertains dans des contextes de soins complexes [88]. Plusieurs approches basées sur l'apprentissage automatique ont été développées pour aider à la prise de décision personnalisée à différentes étapes, du diagnostic, l'identification de la cause de l'état du patient, au pronostic individualisé et aux thérapies ciblées. Par exemple, les techniques d'apprentissage automatique se sont avérées utiles pour développer des scores de risque polygénique qui peuvent aider à identifier les individus présentant un risque génétique élevé de maladie [89]. Les modèles de ML ont été utilisés comme outil de soutien pour identifier des maladies et des types de tumeurs spécifiques à partir de l'analyse de données cliniques/pathologiques [88]. Le ML s'est également avéré utile pour comprendre comment les médicaments et les composés peuvent avoir un impact sur les caractéristiques d'une cellule [90].

### **II.3.2 Adoption de l'apprentissage automatique dans les services de santé**

L'adoption des approches de ML dans les soins de santé est confrontée à des défis complexes liés à deux problèmes principaux [17] : des préoccupations techniques concernant

entre autres la qualité des données ou le problème de la « boîte noire », et des préoccupations humaines concernant l'éthique, le juridique et l'acceptabilité liés au ML :

**Qualité et provenance des données** - Le processus d'apprentissage automatique nécessite de grandes quantités de données pour entraîner les algorithmes. La qualité de ces données d'entraînement est corrélée à la qualité des résultats à obtenir. Si les données sont incomplètes ou ne reflètent qu'une partie de la population, le modèle résultant ne sera pertinent que pour le groupe représenté dans l'ensemble de données. Cela peut avoir des effets négatifs imprévus si le modèle est utilisé pour des patients qui ne font pas partie de ce groupe. Des algorithmes entraînés sur des données générées par certains professionnels de santé sont susceptibles de produire des biais liés à la variabilité des pratiques cliniques et d'impliquer des modèles qui ne reflètent pas nécessairement la situation clinique des patients [15]. En effet, il est fréquent que les cliniciens ne saisissent pas toutes les informations relatives aux patients pour des raisons de confidentialité ou de manque de temps, ou en raison de phénomènes difficiles à saisir, tels que des problèmes psychologiques ou des déterminants sociaux [91]. Cela pose la question de la provenance des données et représente des problèmes liés aux biais des algorithmes utilisés pour la prise de décision [92]. Il faut mettre l'accent sur la transparence, la simplification, la compréhension et l'utilité clinique des ensembles de données, et s'interroger sur l'applicabilité clinique de chaque résultat.

**Problème de la « boîte noire »** - Les algorithmes de ML peuvent fournir des modèles de prédiction fiables si la qualité des données utilisées est appropriée et adaptée au problème. Cependant, ces modèles sont compliqués à comprendre pour un médecin, qui doit savoir comment un modèle traduit les données d'entrée en décisions de sortie. Ce problème de « boîte noire » peut entraîner un manque de confiance ou un sentiment d'inquiétude et de controverse quant à la fiabilité du ML. Ces lacunes dans la connaissance du modèle lui-même et la compréhension du processus d'apprentissage [9] exigent que les résultats des algorithmes soient clairement inclus dans le processus décisionnel de la pratique médicale [5]. En outre, une plus grande importance devrait être accordée à l'explicabilité des résultats obtenus par les modèles de ML. Il s'agit également d'une exigence légale [93].

**Préoccupations humaines** - L'incapacité des prédictions à être transparentes, claires ou dignes de confiance par les médecins, affecte négativement leur utilisation. Les risques

potentiels associés aux données et leur effet sur la sécurité des patients, et la crainte que les médecins puissent être remplacés par des systèmes de ML ont créé des obstacles à l'intégration du ML dans la pratique clinique. Les patients et le public en général font davantage confiance aux médecins qu'à l'IA médicale [94]. Les médecins sont irremplaçables et devraient toujours prendre la décision clinique finale. Pour que leurs patients bénéficient des progrès de l'IA, les médecins doivent avoir un contrôle direct et une vision concrète des recommandations d'un système de ML. Ils doivent être en mesure d'interagir facilement avec les outils d'aide à la décision et d'interpréter les résultats générés par les algorithmes de ML [7].

Le concept d'intelligence hybride a émergé pour surmonter les limites de l'IA [95]. L'hybridation de l'intelligence humaine et de l'intelligence artificielle est une nouvelle forme de collaboration humain-machine destinée à améliorer l'un et l'autre et à faire ressortir le plus grand potentiel des deux [96]. Dans cette optique, les humains peuvent assister le processus de ML de sorte que l'IA puisse aider à la prise de décision humaine. Dans leur étude sur la prédiction hybride, L. Hong et al. [97] ont utilisé une analyse théorique basée sur des preuves empiriques pour démontrer comment les humains et les algorithmes peuvent coopérer pour faire de meilleures prédictions. L'étude a montré que l'inclusion des humains offre plus de précision en réduisant les erreurs, grâce à leur capacité à utiliser des données plus concrètes. Cela a notamment pour avantage d'accroître la confiance dans les prédictions hybrides et, de manière contre-intuitive, dans les algorithmes. L'interaction hybride humain-algorithme avec les capacités de perception et de cognition humaine a le potentiel d'améliorer l'adaptabilité et la précision de l'IA dans le traitement de tâches complexes [98]. Une telle perspective pourrait être révolutionnaire dans le domaine médical.

En guise de conclusion, les principaux critères susceptibles de limiter les inquiétudes quant à l'adoption des systèmes de ML dans les services de santé sont :

1. La transparence, la compréhensibilité et l'utilisation de données approuvées pour leur utilité clinique [99].
2. L'implication des médecins dans le processus de développement des modèles de ML [100].
3. La clarté et la précision de la description des performances du modèle [7].

4. La conception des systèmes de prédiction centrée sur l'utilisateur [101].
5. La facilité d'interprétation et l'explicabilité des résultats des modèles de ML grâce à un meilleur dialogue entre le médecin et le système [58].
6. La mise à disposition du médecin d'un contrôle direct et d'une connaissance approfondie des recommandations du système de ML [102].

Les solutions de la MP pourraient tirer pleinement parti des progrès de l'IA. Les algorithmes d'IA peuvent analyser et apprendre plus rapidement à partir de grandes quantités de données sur les patients. Cela permettrait aux systèmes d'aide à la décision clinique basés sur les modèles obtenus de fournir des recommandations personnalisées aux médecins. Toutefois, il est très important que ces recommandations soient fondées sur le raisonnement des médecins et qu'elles restent sous leur contrôle afin de les aider à prendre des décisions cohérentes et sûres.

## CHAPITRE III

---

# SCLÉROSE EN PLAQUES

### Sommaire du chapitre

|   |    |
|---|----|
| III.1 Qu'est-ce que c'est la sclérose en plaques? . . . . .   | 32 |
| III.2 Prévalence . . . . .                                    | 32 |
| III.3 Origines et facteurs de risques . . . . .               | 34 |
| III.4 Formes . . . . .  | 35 |
| III.5 Symptômes . . . . .                                     | 37 |
| III.6 Diagnostic . . . . .                                    | 38 |
| III.7 Pronostic . . . . .                                     | 42 |
| III.8 Traitements et suivi . . . . .                          | 44 |
| III.9 Enjeux et responsabilités vis-à-vis de la SEP . . . . . | 48 |

### III.1 Qu'est-ce que c'est la sclérose en plaques ?

La première description scientifique de la Sclérose En Plaques (SEP) a été mise en évidence en 1868 par un neurologue français, le professeur Jean-Martin Charcot [103]. La SEP est une maladie neurologique inflammatoire auto-immune chronique qui affecte le système nerveux central (SNC). Elle affecte le cerveau, les nerfs optiques et la moelle épinière, qui ont tous en commun d'être constitués de cellules spécifiques : les neurones. L'inflammation provoque des taches appelées plaques ou lésions. Au niveau d'une lésion inflammatoire, la myéline (la gaine protectrice des neurones) disparaît dans un processus appelé démyélinisation [104]. Cette disparition provient des attaques du système immunitaire contre les cellules qui fabriquent la myéline, considérant la gaine de myéline comme un corps étranger déclenchant une réaction inflammatoire. La myéline endommagée peut être remplacée par des cicatrices qualifiées de « sclérose » qui forment des plaques, d'où le nom « sclérose en plaques ».

### III.2 Prévalence

Entre septembre 2019 et mars 2020, la fédération internationale de la sclérose en plaques (MSIF - *Multiple Sclerosis International Federation*) a collecté les données épidémiologiques de 115 pays représentant 87% de la population mondiale [105]. Ces données ont été exploitées pour illustrer les variations géographiques de la prévalence de la SEP en attribuant un score de confiance à chaque pays (voir Figure III. 1) en fonction de 4 critères :

1. La taille de la population couverte.
2. L'année de collecte des données.
3. Le type de source de données.
4. La conformité de la source de données à certains critères méthodologiques (utilisation des critères McDonald 2017, réalisation d'une étape de validation, etc.).

Les scores de confiance ont été attribués en fonction des scores totaux de tous les critères. De manière générale, l'étude a montré qu'un total approximatif de 2.8 millions de personnes vivent avec la SEP dans le monde (35.9 pour 100 000 habitants).

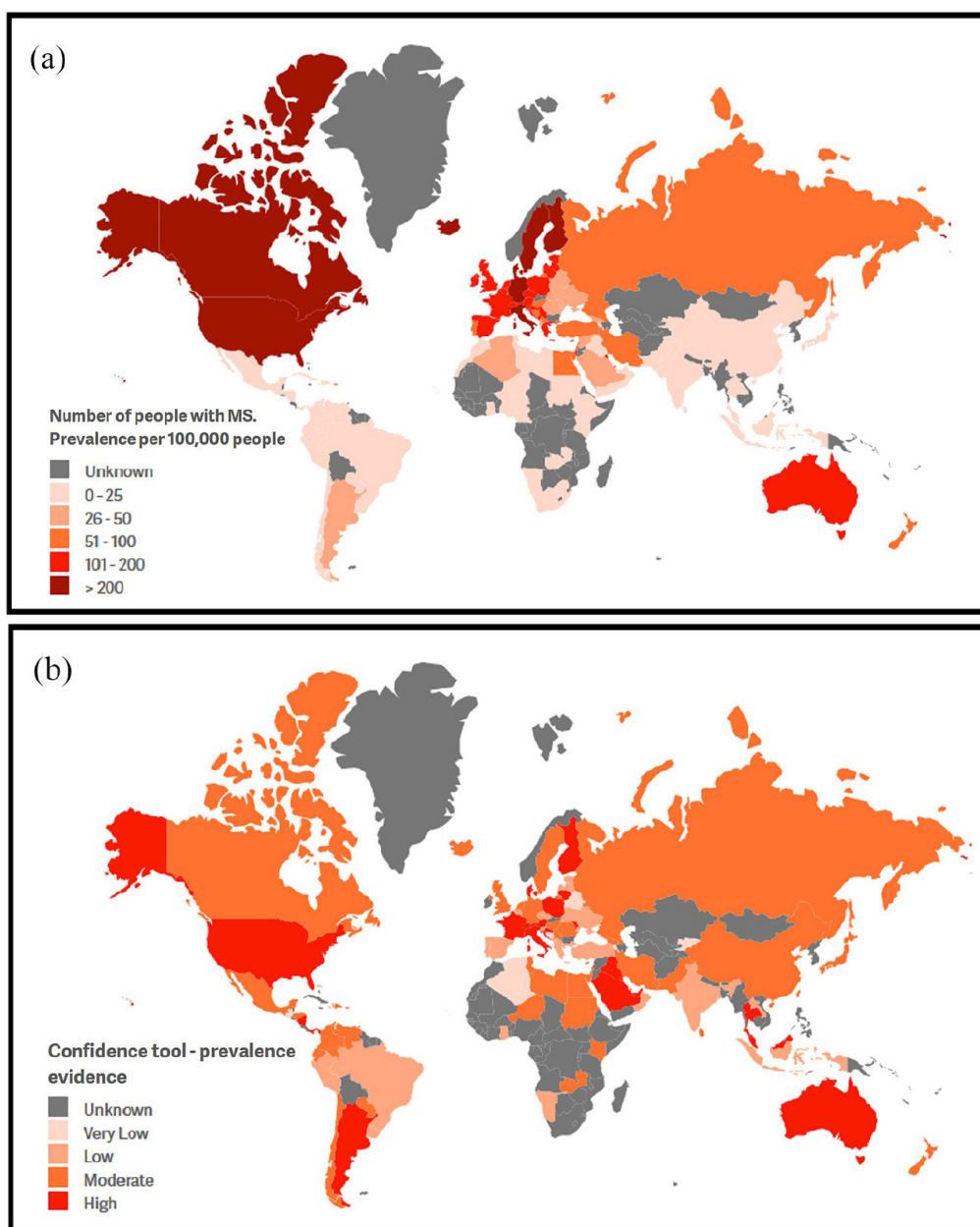


FIGURE III.1 – Cartes montrant la variation géographique de la prévalence de la SEP et des scores de confiance des données par pays (source : C. Walton et al. [105])

(a) Représente le nombre de personnes atteintes de sclérose en plaques (prévalence pour 100 000 personnes). (b) Représente le score de confiance des preuves de prévalence par pays.

Conformément à la Figure III.1, la France est une zone géographique à forte prévalence de la SEP. La dernière estimation en France datant du 31 décembre 2012 a identifié 99 123 personnes atteintes de SEP, correspondant à un taux de prévalence brut global de 151,2 pour 100 000 habitants [106].



### III.3 Origines et facteurs de risques

Les causes et l'origine de la SEP restent inconnues [107], il existe cependant diverses hypothèses sur les facteurs de risque.

**Sexe** - Les femmes représentent environ trois quarts des personnes atteintes de la SEP pour des raisons inconnues [108].

**Prédisposition génétique** - La probabilité de développer la SEP est plus élevée dans les familles déjà touchées par la maladie que dans la population générale. Cependant, la SEP ne présente pas un caractère héréditaire clair. Plusieurs gènes (dont le gène HLA DRB1) peuvent entraîner une augmentation du risque de développer la SEP [109].

**Environnement** - La SEP est présente dans toutes les régions du monde, mais sa prévalence varie considérablement et est toujours plus élevée dans l'hémisphère nord (États-Unis, Canada et pays scandinaves), dans le sud de l'Australie et en Nouvelle-Zélande [109]. La répartition des cas de SEP a conduit les chercheurs à supposer que celle-ci a un lien avec la vitamine D. En 2004, une étude a révélé que les femmes qui prenaient un supplément quotidien de vitamine D par rapport aux femmes qui n'en prenaient pas, réduisaient leur risque de développer une sclérose en plaques de 40% [110]. Les données actuelles ne permettent malheureusement pas de déterminer si une supplémentation en vitamine D peut affecter l'évolution de la maladie chez l'homme.

**Infections** - Certaines infections ou maladies contractées pendant l'enfance peuvent être impliquées. Le virus d'Epstein-Barr (EBV - *Epstein-Barr Virus*) [111], qui se manifeste par une infection des voies respiratoires ou par une mononucléose infectieuse, a été identifié dans plusieurs études comme un facteur de développement de la SEP. Il a été signalé que le risque de SEP augmente après une infection par l'EBV [112] et que le fait d'avoir une mononucléose infectieuse augmente le risque de développer une SEP [113].

**Mode de vie** - Il a été démontré que le tabagisme joue un rôle important dans la détermination de la susceptibilité à la SEP, cependant son effet sur la progression de la maladie est moins certain [114]. Il a également été révélé que l'obésité pendant l'enfance

et l'adolescence augmente le risque de SEP [115]. D'autres facteurs de risque ont été repérés, tels que l'implication du sel alimentaire dans la pathogenèse de l'auto-immunité du SNC, la dose de consommation d'alcool, le stress psychologique [116]. Hors, ces facteurs n'ont pas été mis en évidence dans des études cliniques.

De manière générale, on ne sait pas vraiment si les facteurs de risque de la SEP agissent en séquence et dépendent les uns des autres, ou s'ils agissent indépendamment et de manière additive ou multiplicative. Étant donné qu'aucun facteur ne s'est encore avéré être présent chez tous les patients atteints de SEP, les voies causales sont susceptibles de différer d'un individu à l'autre [116].

### III.4 Formes

L'âge moyen de survenue de la SEP se situe généralement entre 20 et 40 ans [117]. Or, il existe des formes pédiatriques (âge d'apparition avant 18 ans) et des formes qui débutent après 40 ans. En 1996, une classification des formes de SEP a été proposée suite à un consensus entre des médecins internationaux de la SEP pour concevoir correctement les essais cliniques, assurer l'homogénéité des patients recrutés, faciliter la communication entre les médecins, et aider à déterminer qui peut bénéficier de l'utilisation du médicament [118]. Cette classification est rapidement devenue une partie essentielle de la pratique clinique et de la recherche. Elle a joué un rôle déterminant dans le processus d'approbation des nouveaux traitements de la SEP. Cependant, au fil du développement des connaissances sur la pathologie de la SEP, les limites des formes purement cliniques, dépourvues d'imagerie et de corrélations biologiques, sont devenues évidentes. Un autre consensus a été conclu en 2012 [119] pour revoir les formes cliniques afin de fournir une terminologie améliorée tout en incorporant l'imagerie, les biomarqueurs, etc. Les formes définies en 1996 ont été conservées tout en améliorant la catégorisation par l'introduction de nouveaux descripteurs de l'activité et de la progression de la maladie, le syndrome radiologiquement isolé et le syndrome cliniquement isolé [120].

**Le Syndrome Cliniquement Isolé (SCI)** - Il se réfère à un premier épisode de démyélinisation inflammatoire dans le système nerveux central qui pourrait devenir une SEP si une activité supplémentaire se produit.

**Le Syndrome Radiologiquement Isolé (SRI)** - C'est une situation compliquée où des résultats d'imagerie suggèrent une démyélinisation inflammatoire en l'absence de signes ou de symptômes cliniques. Le SRI n'a pas été considéré comme une forme de SEP en soi, car il n'y a pas de preuve clinique de démyélinisation (un critère actuel pour le diagnostic de la SEP) et les résultats d'IRM seuls peuvent être non spécifiques. Cependant, le SRI peut faire suspecter une SEP, selon la morphologie et la localisation des lésions IRM détectées.

**La Sclérose En Plaques Récurrente-Rémittente (SEP-RR)** - C'est la forme la plus fréquente de la maladie, environ 85% des patients atteints de SEP ont d'abord une SEP-RR. Elle débute en moyenne vers l'âge de 30 ans et touche plus les femmes que les hommes, avec un ratio de trois femmes pour un homme [121]. Elle se caractérise par des épisodes d'aggravation aiguë du fonctionnement neurologique (nouveaux symptômes ou aggravation des symptômes existants) avec récupération totale ou partielle et aucune progression apparente de la maladie. Une SEP-RR active présente des signes de nouvelles poussées (apparition de nouveaux symptômes ou l'aggravation de symptômes déjà existants, pendant plus de 24 heures, en dehors d'une période de fièvre), de nouvelles lésions renforcées par le gadolinium et/ou de nouvelles lésions T2 ou d'un élargissement des lésions T2 à l'IRM sur une période donnée. Elle est inactive lorsqu'elle ne présente aucun signe d'activité de la maladie.

**La Sclérose En Plaques Primaire Progressive (SEP-PP)** - C'est la forme la plus rare de la SEP. Elle touche environ 15% des patients de SEP et évolue rapidement. L'âge moyen de début des symptômes est d'environ 40 ans et touche autant les hommes que les femmes [121]. Elle se caractérise par une aggravation constante de la fonction neurologique depuis le début des symptômes, sans poussées ni rémissions initiales. Des poussées peuvent parfois survenir chez certains patients.

**La Sclérose En Plaques Progressive Secondaire (SEP-PS)** - Elle est caractérisée par une phase initiale récurrente rémittente suivie d'une forme progressive avec ou sans poussées occasionnelles. Actuellement, il n'existe pas de critères cliniques, d'imagerie, immunologiques ou pathologiques clairs pour déterminer le point de transition entre la SEP-RR et la SEP-PS. La transition est généralement déterminée par

l'aggravation progressive des symptômes neurologiques.

### III.5 Symptômes

Les symptômes de la SEP sont variables selon les zones du SNC affectées. Ils peuvent résulter de l'atteinte des voies sensorielles, motrices, visuelles et du tronc cérébral [109]. Les premières manifestations de la SEP peuvent concerner :

**Des troubles de la vision** - Les troubles ophtalmologiques sont révélateurs de la maladie. Ils sont le plus souvent caractérisés par une vision floue, double, perte de vision complète ou partielle, habituellement d'un œil à la fois, douleurs lorsqu'on bouge, les douleurs oculaires, mouvements oculaires involontaires, « voile » devant un œil, paupière qui tremble, etc. Ces troubles sont causés par une névrite optique (une atteinte du nerf optique).

**Des troubles sensitifs** - Ces troubles peuvent toucher différentes parties du corps (membres, thorax, visage...). Le signe de Lhermitte qui est une sensation de décharge électrique dans le dos déclenché par une flexion de la tête est très évocateur de la maladie. Le patient peut également ressentir des fourmillements, des démangeaisons, une diminution de la sensibilité du toucher, une faiblesse musculaire, avoir du mal à reconnaître le chaud et le froid, etc.

**Des troubles de la motricité** - Les symptômes moteurs sont fréquents au début de la maladie. Ils peuvent se traduire par une fatigabilité anormale à la marche, une atteinte de la mobilité de toute une partie du corps, des réflexes perturbés, etc. Les troubles de l'équilibre sont fréquents, ils sont à l'origine de nombreuses chutes. Cela s'explique par une atteinte de certaines régions du SNC responsables de la coordination des mouvements et de l'équilibre.

**Des troubles neurologiques (symptômes transitoires)** - Ils se manifestent notamment par un trouble de l'articulation. Cela est dû à une atteinte des nerfs qui assurent la motricité des muscles des cordes vocales. Les troubles neurologiques peuvent également se traduire par une maladresse gestuelle.

Outre les signes neurologiques, **des signes généraux** sont souvent présents au début de la maladie ou lorsque celle-ci évolue. Ils se manifestent notamment par des troubles

urinaires, des troubles du sommeil, de la fatigue, des troubles sexuels, des troubles de l'humeur, des troubles cognitifs, des troubles intestinaux, etc.

Il faut noter que ces symptômes peuvent être liés à d'autres maladies. Leur apparition n'indique pas systématiquement la présence d'une SEP.

### III.6 Diagnostic

L'hétérogénéité de la SEP en termes de symptômes, dont certains peuvent être spontanément résorbés et d'autres attribués à d'autres pathologies, a rendu nécessaire la combinaison de preuves cliniques, biologiques et radiologiques et le suivi de l'évolution des symptômes pour poser le diagnostic. Ce dernier ne peut être établi que par la démonstration clinique et/ou radiologique de lésions dans le SNC qui sont disséminées dans l'espace et dans le temps [109]. Les examens paracliniques suivants peuvent fournir des preuves diagnostiques :

**L'Imagerie par Résonance Magnétique (IRM)** - Indicateur très important pour confirmer le diagnostic de la SEP en produisant des images très détaillées du cerveau et de la moelle épinière. Ces images illustrent les lésions sous la forme de signaux de la substance blanche (responsable de la propagation de l'information dans le système nerveux) et permettent d'exclure d'autres maladies. La visualisation des lésions sur l'IRM peut se faire de différentes manières en fonction des séquences IRM. La séquence la plus utilisée dans le cerveau est la séquence FLAIR (*Fluid Attenuated Inversion Recovery*) [122] pondérée en T2. Les séquences pondérées en T2 permettent de visualiser facilement les lésions qui apparaissent en hypersignal, notamment les lésions périventriculaires et juxtacorticales ou corticales et les lésions de la moelle épinière (lésions médullaires) (voir Figure III.2). La séquence T1 peut également être réalisée en injectant un produit de contraste, le gadolinium, qui permet de fournir des informations sur le caractère récent ou ancien des lésions. Dans le cas d'une lésion récente « active », le gadolinium peut pénétrer dans le cerveau, révélant une rupture de la barrière hémato-encéphalique (barrière isolant le cerveau du reste du corps). Les lésions les plus anciennes ne sont pas rehaussées par le gadolinium et sont donc dites « non actives ».

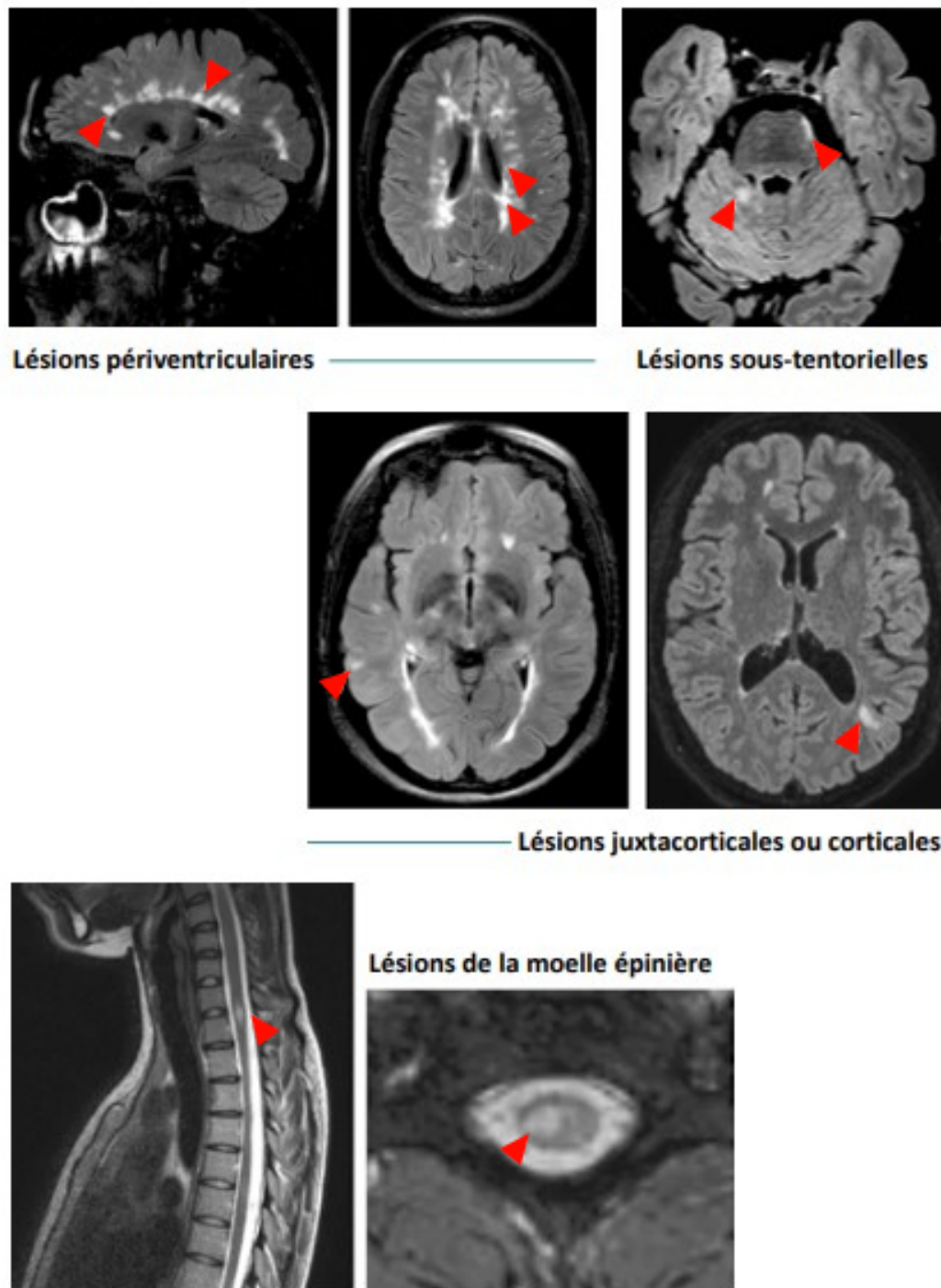


FIGURE III.2 – Quatre localisations préférentielles pour les lésions inflammatoires (source : ARSEP – Le point sur la SEP [121])

*Les zones marquées en rouge sont les lésions.*

**Examen électro-physiologique** - Étude des potentiels évoqués permettant d'identifier des lésions cliniquement silencieuses dans les voies visuelles, le tronc cérébral et

la moelle épinière. Il s'agit de mesurer le temps nécessaire au cerveau pour recevoir et interpréter un signal visuel, auditif ou sensoriel. Les enregistrements sont effectués à l'aide d'électrodes placées sur la tête du patient pour mesurer la réponse électrique du cerveau aux stimuli. La présence de plaques de démyélinisation entraîne un ralentissement de l'influx nerveux.

**Ponction lombaire** - Consiste à prélever un échantillon du liquide dans lequel baignent le cerveau et la moelle épinière, le Liquide Céphalo-Rachidien (LCR), par ponction à l'aide d'une fine aiguille dans la région inférieure de la colonne vertébrale. L'objectif est de rechercher des marqueurs inflammatoires en analysant ce liquide. Les marqueurs inflammatoires du LCR sont présents jusqu'à 85% chez des patients atteints de SEP. La ponction lombaire n'est pas obligatoire lorsque les critères diagnostiques cliniques et/ou IRM sont remplis et confirmés. Elle reste cependant fortement recommandée, car elle permet de confirmer la présence d'une inflammation au sein du LCR par la mise en évidence de Bandes Oligoclonales (BOC), qui correspondent à la production d'anticorps.

Une réévaluation des critères de diagnostic de la maladie se fait au fil de l'évolution des connaissances scientifiques. La dernière révision correspond aux critères de McDonald 2017 [123] (voir Tableau III.1). Le diagnostic des formes récurrentes se manifeste par la présence de deux poussées avec des symptômes différents et à des moments différents, ce qui correspond à une dissémination dans l'espace (DIS - *Dissemination In Space*) et une dissémination dans le temps (DIT - *Dissemination In Time*) et répond donc aux critères diagnostiques de la SEP. L'IRM permettra alors de confirmer le diagnostic. D'autres examens peuvent être recommandés. Par contre, si le patient présente une première et unique poussée clinique, les critères de DIS et DIT seront alors validés par une IRM et/ou une ponction lombaire.

Le diagnostic de la forme progressive primaire est établi lorsque le patient présente des symptômes neurologiques s'aggravant progressivement depuis au moins un an, en dehors de toute poussée clinique. Des examens complémentaires sont indispensables pour confirmer le diagnostic et éliminer les diagnostics différentiels [121].

TABLEAU III.1 – Critères diagnostiques de McDonald 2017.

| <b>Formes</b>       | <b>Dissémination dans l'espace</b>   | <b>Dissémination dans le temps</b>   |
|---------------------|--|--|
| Formes rémittentes  | <ul style="list-style-type: none"> <li>- Survenue de poussées dans différents endroits, avec des symptômes différents</li> <li>- Présence d'au moins une lésion T2 dans au moins deux des 4 localisations du SNC : périventriculaire, juxtacorticale ou corticale, sous-tentorielle et médullaire</li> </ul>   | <ul style="list-style-type: none"> <li>- Survenues de poussées à des moments différents</li> <li>- Présence simultanée de lésions actives et non actives quel que soit le moment de l'IRM</li> <li>- Nouvelle(s) lésion(s) T2 et/ou gadolinium positives sur une IRM de suivi quel que soit le moment de l'IRM de référence et de l'IRM de suivi</li> <li>- Ponction lombaire : présence de BOC dans le LCR</li> </ul> |
| Formes progressives | <ul style="list-style-type: none"> <li>- Aggravation progressive des symptômes neurologiques depuis plus d'un an</li> <li>- Présence d'au moins 2 des 3 critères suivants :                             <ol style="list-style-type: none"> <li>1. Au moins une lésion T2 dans une ou plus des 3 localisations : périventriculaire, juxtacorticale ou corticale, sous-tentorielle</li> <li>2. Au moins 2 lésions médullaires</li> <li>3. Présence de BOC dans le LCR</li> </ol> </li> </ul> |  |



### III.7 Pronostic

L'histoire naturelle de la SEP et l'extrême variabilité de son évolution ont rendu compliqués le suivi, le pronostic et les décisions thérapeutiques. Deux nouvelles notions de la SEP ont émergé pour mieux guider les décisions thérapeutiques et évaluer le pronostic :

- **L'activité** : définie par la présence de poussées et/ou l'apparition de nouvelles lésions sur l'IRM.
- **La progression** : déterminée par la progression du handicap et/ou des signes neurologiques.

Le degré de handicap des patients atteints de SEP est quantifié par une échelle de cotation appelée EDSS (*Extended Disability Scale Score*). L'EDSS quantifie le handicap dans huit systèmes fonctionnels (SF) : pyramidal, cérébelleux, sensitif, tronc cérébral, sphincters, vision, mental et autres [124]. Un score numérique de sévérité croissante (0 à 6 ou 7) est attribué à chaque SF. Le score global de l'échelle est mesuré sur une échelle de 20 niveaux, de 0 à 10 par demi-points. Une échelle EDSS 0 correspond à une absence de handicap, une échelle EDSS entre 1 et 3 correspond à un retentissement modéré et à une situation de faible handicap, une échelle EDSS entre 3.5 et 5 correspond à un handicap ne nécessitant pas l'utilisation systématique d'une aide à la marche, une échelle EDSS entre 5.5 et 7 correspond à la nécessité d'une aide technique, et une échelle EDSS entre 7.5 et 9 correspond à la nécessité de l'utilisation permanente d'un fauteuil roulant. Une étude réalisée par E. Leray et al. [125] s'est concentrée sur la progression du handicap dans la SEP. Sur la base des données de 2290 patients atteints de SEP, ils ont comparé la progression de l'invalidité au cours deux phases sur la base de l'échelle EDSS simplifiée (voir Figure III.3). La première phase représente la maladie depuis le début de la SEP jusqu'à l'atteinte d'un EDSS 3 irréversible (EDSS persistant au moins 6 mois). La deuxième phase commence à partir du niveau d'invalidité 3 jusqu'à atteindre l'EDSS 6. Les résultats ont montré que la progression du handicap en phase 2 est indépendante de sa progression en phase 1. Par ailleurs, la durée de la phase 2 reste remarquablement constante.

Peu importe la forme progressive de la SEP, elle peut être active ou non. Dans le cas d'une forme progressive, l'aggravation n'est pas constante, l'état du patient peut rester stable durant plusieurs années. Il a été démontré que plusieurs facteurs peuvent être pré-

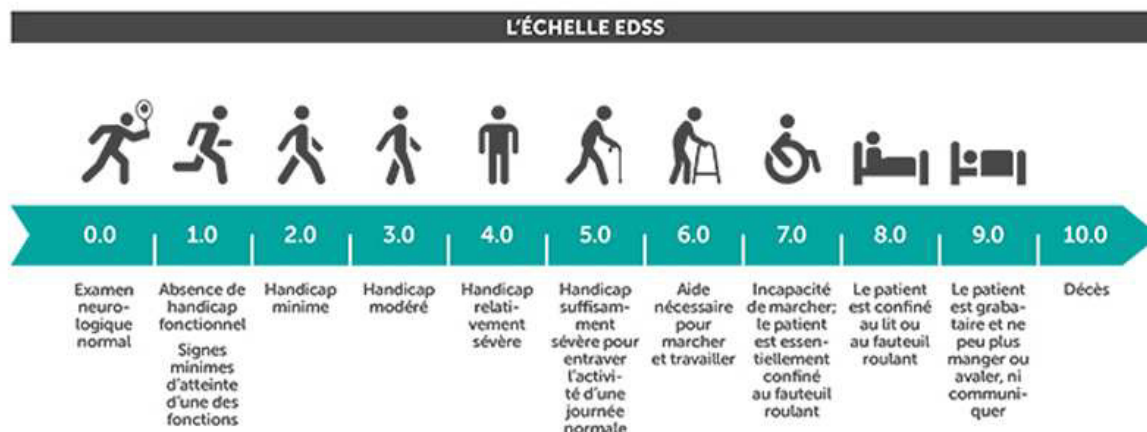


FIGURE III.3 – Échelle EDSS-simplifiée (source : sep.g-station.com)

dictifs de la progression de la SEP, tels que l'âge, le sexe, l'origine ethnique, la topologie des poussées présentées et leur fréquence, l'accumulation du handicap après le début de la maladie, les lésions sur l'IRM de base, la présence de bandes oligoclonales spécifiques dans le LCR et la présence de lésions dans la moelle épinière. D'autres biomarqueurs de laboratoire tels que les neurofilaments (filament intermédiaire des neurones formé de protéines spécifiques) et d'autres mesures d'imagerie plus avancées telles que l'atrophie du cerveau entier, le volume de la matière grise du cerveau, le volume de la matière grise de la moelle épinière, sont en cours de développement pour être disponibles pour une utilisation clinique afin d'améliorer le pronostic et le suivi thérapeutique de la SEP [117].

À un stade précoce de la SEP, plusieurs facteurs pronostiques peuvent être évalués pour prédire l'évolution de la maladie. Parmi les principaux facteurs pris en compte, on trouve la forme initiale de la maladie (SEP-RR, SEP-PP ou SEP-PS), la présentation clinique initiale et la sévérité des premiers symptômes, l'âge au moment du diagnostic [126], le niveau de handicap précoce mesuré par des échelles telles que l'EDSS, l'activité inflammatoire évaluée par l'IRM ou des marqueurs spécifiques, et la réponse au traitement [127]. Il est important de souligner que l'évaluation des facteurs pronostiques est complexe et les résultats peuvent varier d'une personne à l'autre. Il est crucial de suivre régulièrement l'évolution de la maladie et de réévaluer les facteurs pronostiques afin d'adapter le traitement et la prise en charge.

### III.8 Traitements et suivi

À ce jour, aucun traitement ne permet de guérir la SEP. Cependant, les traitements sont de plus en plus efficaces pour prévenir les poussées et l'apparition d'un handicap permanent. On peut distinguer deux types de traitement de la SEP : les traitements modificateurs de la maladie, qui sont généralement spécifiques à la SEP, et les traitements symptomatiques, qui sont souvent utilisés dans différents contextes pathologiques [128]. Les traitements de fond visent à ralentir l'évolution de la SEP. Ils agissent sur les mécanismes responsables des phénomènes inflammatoires et immunitaires. Il existe 3 types de traitements modificateurs de l'évolution de la SEP [121] (voir tableau III.2) :

**Traitements de première ligne** - Ils représentent le premier traitement (depuis 1990). Ils comprennent les traitements injectables et les comprimés. Généralement, les traitements de ce type ont une réduction similaire ou proche de 30% du risque de poussée par rapport au placebo.

**Traitements de seconde ligne** - Utilisés pour les formes cliniquement / radiologiquement sévères de la maladie ou lorsque les traitements de première ligne ne permettent pas de contrôler la maladie (progression des poussées ou des lésions).

**Traitements de troisième ligne** - Utilisés pour les formes agressives ou lorsque les traitements de deuxième ligne ne sont pas suffisants. Ils ont une forte efficacité (prolongée après leur arrêt) mais ils peuvent avoir de forts effets toxiques. Ils ne sont donc utilisés que pour une durée limitée.

TABEAU III.2 – Traitements de fond de la SEP (source : ARSEP – Le point sur la SEP [121])

| Type           | Traitement             | Administration                                   | Rythme               | Effets indésirables   | Grossesse      |
|----------------|------------------------|--|----------------------|---|----------------|
| Première ligne | Interférons            | Injection  | Tous les 2 jours à   | Syndrome pseudogrippal  | Possible       |
|                |                        | sous-cutanée ou intramusculaire (auto-injection) | tous les 15 jours    | Effets sur le foie : surveillance biologique par prise de sang  |                |
|                | Acétate de glatiramère | Injection  | - 1 fois par jour    | Nodules sous-cutanés,   | Possible       |
|                |                        | sous-cutanée (auto-injection)                    | - 3 fois par semaine | « placards rouges » cutanés, nécrose au point d'injection<br>Malaises : impression d'oppression thoracique avec gêne respiratoire                               |                |
|                | Tériflunomide          | Voie orale                                       | 1 fois par jour      | Affinement du cheveu<br>Diarrhées, nausées, coliques<br>Augmentation de la tension artérielle<br>Effets sur le foie : surveillance biologique par prise de sang | Contre-indiqué |

TABLEAU III.2 - Traitements de fond de la SEP (source : ARSEP – Le point sur la SEP [121]) (Suite)

| Type           | Traitement        | Administration | Rythme  | Effets indésirables   | Grossesse      |
|----------------|-------------------|----------------|---|---|----------------|
|                | Diméthyl-fumarate | Voie orale     | 2 fois par jour   | Troubles digestifs à type de diarrhées, nausées, douleurs abdominales<br>Douleurs abdominales hautes<br>Bouffées vasomotrices<br>Lymphopénies : surveillance biologique de la numération sanguine | Contre-indiqué |
| Deuxième ligne | Fingolimod        | Voie orale     | 1 cp/j à 0,5 mg<br>1 cp/j à 0,25 mg chez l'enfant de moins de 50 kg (10-18 ans) | Troubles du rythme cardiaque<br>Œdème maculaire<br>Cancer cutané (basocellulaire)<br>Recrudescence de zona (réactivation du virus varicelle-zona), herpès et HPV (papillomavirus humain)          | Contre-indiqué |
|                | Cladribine        | Voie orale     | Cycles de plusieurs jours   | Infections<br>Recrudescence de zona et herpès   | Contre-indiqué |

TABLEAU III.2 - Traitements de fond de la SEP (source : ARSEP – Le point sur la SEP [121]) (Suite)

| Type                   | Traitement   | Administration     | Rythme   | Effets indésirables   | Grossesse      |
|------------------------|--------------|--------------------|--|---|----------------|
|                        | Natalizumab  | Voie intraveineuse | 1 fois toutes les 4 semaines en hospitalisation                              | Leucoencéphalopathie multifocale progressive  | Possible       |
|                        | Ocrélizumab  | Voie intraveineuse | 1 fois tous les 6 mois en hospitalisation                                    | Infections des voies aériennes supérieures<br>Recrudescence de zona et herpès   | Contre-indiqué |
| <b>Troisième ligne</b> | Mitoxantrone | Voie intraveineuse | 1 fois par mois pendant 6 mois en hospitalisation                            | Troubles digestifs à type de nausées, vomissements<br>Insuffisance cardiaque<br>Leucémies<br>Infections                 | Contre-indiqué |
|                        | Alemtuzumab  | Voie intraveineuse | 1 cycle de 5 jours la première année<br>1 cycle de 3 jours la deuxième année | Purpura thrombopénique<br>auto-immun Hyper et/ou hypothyroïdie<br>Atteinte rénale (glomérulonéphrite extra-membraneuse) | Contre-indiqué |

Les traitements symptomatiques visent à améliorer la qualité de vie des patients en prenant en charge les symptômes qui altèrent la vie quotidienne, notamment la fatigue, la douleur, la dépression, les troubles de l'équilibre, de la cognition, urinaires, sexuels, du transit, du sommeil, etc. Certains facteurs étant liés à d'autres comme les troubles de la marche, de la vision, de la coordination, etc. font que les thérapies sont proposées en fonction d'une analyse précise de la situation. Des médicaments et des activités physiques (par exemple, la kinésithérapie) sont prescrits par les médecins pour la gestion de ces symptômes afin de maintenir la vie des patients. Le suivi des patients traités ou non est essentiel dans la SEP. Une IRM cérébrale doit être réalisée régulièrement au début du traitement ainsi qu'un suivi clinique pour surveiller l'activité et la progression de la maladie. Le type de traitement peut être modifié en fonction de son efficacité vis-à-vis de l'activité de la maladie.

### **III.9 Enjeux et responsabilités vis-à-vis de la SEP**

La SEP a non seulement des conséquences négatives pour le patient individuel, mais aussi pour la société. Les patients atteints de SEP ont un taux de consultations et de consommation de médicaments nettement plus élevé et un taux d'emploi très faible, ce qui entraîne un coût socio-économique élevé [129]. Le sexe, la famille, la culture, la société, etc. sont des facteurs socioculturels qui influencent le comportement des patients atteints de SEP [130]. Dans certaines cultures où l'homme de la famille est censé être le principal responsable économique et le décideur de la famille, les troubles de mobilité liés à la SEP peuvent causer différents problèmes. Dans cette dynamique familiale, l'homme atteint de SEP peut s'efforcer et se battre pour continuer à travailler malgré ses problèmes de mobilité. Si la personne est trop fatiguée, risque de se mettre en danger en continuant à travailler. Dans ce groupe culturel, on attend des femmes qu'elles soient soumises et qu'on s'occupe d'elles, par conséquent la situation de dépendre des hommes devient plus problématique pour les femmes atteintes de SEP. La personne à mobilité réduite n'est pas seulement confrontée à un défi physique, mais doit également surmonter des attitudes et limites sociales afin de participer pleinement à la vie de la communauté. Les personnes à faible revenu ou disposant de peu de ressources financières sont fortement désavantagées lorsqu'il s'agit d'obtenir des dispositifs de mobilité appropriés et/ou de modifier

leur environnement immédiat en fonction des besoins.

En comparaison avec la population générale, les patients atteints de SEP présentent un taux de suicide plus élevé. Ce phénomène est associé à plusieurs facteurs de risque : gravité de la dépression, isolement social, revenus plus faibles, progression précoce de la maladie, etc. [131]. Lorsque le patient éprouve des besoins existentiels, spirituels ou religieux en rapport avec sa maladie et ses conséquences, pour son bien-être, il peut être intéressant que les médecins abordent les aspects psychosociaux en complément des perspectives biomédicales [132]. L'ouverture aux préoccupations existentielles des patients souffrant de douleur pourrait être un aspect important des soins apportés à la douleur chronique [133], mais aussi un grand défi pour les médecins qui ne partagent pas les mêmes croyances et/ou convictions avec le patient.

Les enjeux psychologiques, sociaux et psychiatriques jouent un rôle majeur dans la qualité de vie des patients atteints de SEP. La famille et/ou les amis peuvent apporter un soutien précieux au patient en lui offrant un environnement sociable et solidaire qui favorise son dynamisme dans la vie. Quant aux services de santé, ils peuvent jouer un rôle très important en facilitant les soins et le soutien psychologique (dès le diagnostic) pour les patients pour qui cela n'est pas possible [134]. Ceci en assurant une thérapie et un suivi pour les patients à faibles revenus, en facilitant les déplacements des médecins, en améliorant l'interaction médecin-patient grâce à des outils d'aide à la communication, etc.

La SEP est une maladie multidimensionnelle et complexe. Les patients atteints de SEP ont des profils très différents et beaucoup de traitements sont disponibles (et donc de données). D'où l'intérêt d'un accompagnement du numérique capable d'intégrer ces dimensions pour mieux assister les praticiens.





# PROJECTION DE DONNÉES POUR L'AIDE À LA DÉCISION DANS LA SEP

## Sommaire du chapitre

|  |    |
|--|----|
| Introduction . . . . .   | 52 |
| IV.1 État de l'art . . . . .   | 52 |
| IV.2 Hypothèse . . . . .   | 53 |
| IV.3 Enjeux des données . . . . .  | 53 |
| IV.4 Outil de médecine de précision pour la SEP : MS Vista . . . . .                               | 54 |
| IV.4.1 Analyse de l'activité . . . . .   | 55 |
| IV.4.2 Conception . . . . .  | 57 |
| IV.4.2.1 Contextualisation : POI / POR . . . . .   | 58 |
| IV.4.2.2 Courbe d'évolution . . . . .  | 58 |
| IV.4.2.3 Projection . . . . .  | 60 |
| IV.4.3 Prototypage . . . . .   | 62 |
| IV.4.3.1 Évaluation en vue d'un développement basé sur l'utilisateur                               | 62 |
| IV.4.3.2 Fonctionnalités de MS Vista . . . . .   | 63 |
| IV.4.4 Évaluation de MS Vista comme outil support de la communication<br>médecin patient . . . . . | 71 |
| Discussion . . . . .   | 74 |

## Introduction

Un des objectifs de la collaboration humain-machine pour l'aide à la décision médicale est de rendre la compréhension des données et leur traitement le plus clair possible pour les personnes qui les utilisent. Ils pourraient permettre de rendre palpables, simples et concrètes des données qui apparaissaient abstraites. Dans ce chapitre, je présente les différentes étapes qui ont permis le développement d'un prototype fonctionnel de système d'aide à la décision clinique (CDSS) pour la SEP : MS Vista. MS Vista est basé sur des concepts de la Médecine de Précision (MP) et le dialogue humain-machine, à travers des interactions avec des données de patients atteints de SEP. Ce travail a été publié dans le journal *Annals of Clinical and Translational Neurology* en tant qu'article de recherche par Ed-Driouch et al. [2].

### IV.1 État de l'art

Plusieurs CDSS ont été développés pour aider à la prise en charge de la SEP, tels que décrits dans l'étude d'Alshamrani et al. [135]. En 2013, Veloso [136] a développé un CDSS pour prédire le handicap à long terme et l'effet du traitement sur les estimations pronostiques en utilisant une cohorte de patients atteints de SEP. En 2014, Gourraud et al. [137] ont développé MS Bioscreen, un CDSS pour iPad permettant de contextualiser les patients atteints de SEP et de visualiser la progression de la maladie en utilisant les données issues d'une cohorte. Pour étendre la portée de MS Bioscreen, Schleimer et al. [138] ont développé Open MS Bioscreen. Celui-ci est conçu pour les patients afin de les aider à prendre des décisions concernant leurs soins. Ces systèmes constituent une aide considérable pour la gestion de la SEP, que ce soit pour l'aide à la décision, la prise en charge des patients ou la prévision de l'effet des traitements. Cependant, ils sont limités par certains problèmes liés à la précision et à la transparence des méthodes et/ou de la qualité des données utilisées. La qualité et l'utilité clinique des données utilisées ne sont pas explicitement présentées par les systèmes. De même, les processus d'aide à la décision sous-jacents n'incluent pas directement les médecins experts de la SEP. L'effet potentiel personnalisé du traitement à suggérer au patient n'est pas tout à fait pris en compte. D'autre part, les outils destinés aux patients pour prendre des décisions concernant leurs soins

peuvent facilement conduire à une confusion d'interprétation. Cela présente un risque pour la santé des patients et pour leur relation avec le médecin.

## IV.2 Hypothèse

La SEP est une maladie complexe qui se caractérise par la variabilité en termes de symptômes qui diffèrent d'un patient à l'autre, en termes de formes, et en termes d'ensemble d'arguments pour poser le pronostic (plus de détails sur la SEP et les facteurs qui y sont liés sont disponibles dans le Chapitre III : Sclérose en plaques). Les patients atteints de SEP nécessitent une prise en charge spécifique et à long terme. Il n'y a pas de consensus sur la prise en charge, qui dépend donc des informations retenues par le médecin. Pour ce faire, les médecins doivent gérer un grand nombre de données tout en faisant intervenir les éléments de connaissance, d'incertitude, d'expérience, de risque et d'intuition. Nous supposons que la collaboration entre les médecins experts et les chercheurs en informatique et en facteurs humains conduira au développement d'un outil de médecine de précision qui aidera les praticiens à interagir facilement avec les données des patients, afin d'obtenir les informations nécessaires pour prendre la meilleure décision au bon moment pour chaque patient.

## IV.3 Enjeux des données

L'utilisation des données est l'une des principales bases de la médecine de précision. Pour évaluer notre hypothèse, nous avons eu accès à deux bases de données de patients atteints de SEP, une cohorte nationale de l'OFSEP (Observatoire Français de la Sclérose en Plaques) et les données d'un essai clinique (NCT00906399) de Biogen, une entreprise internationale de biotechnologie. Ces dernières proviennent d'une étude mondiale multicentrique, randomisée, en double aveugle, en groupe parallèle, contrôlée par placebo sur l'efficacité du peginterféron bêta-1a (traitement anti-inflammatoire) dans la réduction des taux de poussée chez les patients atteints de SEP-RR [139]. Préalablement à l'utilisation des données, nous avons évalué leur qualité, afin d'éviter les conséquences liées à une mauvaise qualité des données.

Dans un premier temps, les données de l'OFSEP étaient par contre difficiles à utili-

ser pour démarrer nos recherches sur l'aide à la décision clinique. Il y avait beaucoup de données manquantes, en particulier les données IRM, qui constituent un élément très important pour la prise de décision dans la SEP. L'exactitude de certaines données a également suscité des inquiétudes, par exemple, on peut trouver des patients qui sont passés d'un score EDSS (mesure de l'incapacité physique) de 6 (handicap sévère) à 0 (pas de handicap) en 6 mois, ce qui peut paraître peu cohérent. Une autre forme de ces données, la cohorte OFSEP HD, est actuellement mise en place pour fournir des données de meilleure qualité. La cohorte OFSEP HD se concentre sur un groupe spécifique de patients sélectionnés dans le réseau OFSEP, dans des centres de soins sélectionnés. L'objectif est d'obtenir une collection complète de données sur les facteurs de la maladie afin de permettre l'évaluation du risque de progression de la maladie dans différentes dimensions [140].

Les données issues de l'essai clinique de BIOGEN sont des données de phase III, ce qui correspond à la phase de confirmation de l'efficacité du traitement et de demande d'autorisation de mise sur le marché. Nous avons exploité le potentiel de la qualité de ces données (précision, accessibilité, cohérence, actualité et exhaustivité) pour développer une preuve de concept d'outil d'aide à la décision basé sur une interaction directe avec les données. C'est la première fois qu'on utilise à la demande des données individuelles de RCT (*Randomized Controlled Trial*).

#### **IV.4 Outil de médecine de précision pour la SEP : MS Vista**

Nous avons collaboré avec des médecins experts de la SEP pour développer MS Vista, prototype fonctionnel de CDSS destiné à soutenir la prise de décision médicale pour les patients atteints de SEP. Pour ce faire, nous nous sommes basés sur la méthodologie de conception centrée sur l'utilisateur des systèmes d'information clinique présentée par Kushniruk [101]. Cette méthodologie vise à optimiser l'efficacité et la convivialité en développant des solutions qui répondent aux exigences et aux attentes des utilisateurs finaux. Cette section représente les différentes phases de développement : analyse de l'activité, conception, prototypage et évaluation.

### IV.4.1 Analyse de l'activité

Nous avons commencé notre analyse de l'activité par des recherches bibliographiques sur la prise de décision, le pronostic et le suivi thérapeutique de la SEP [109, 124, 125, 137, 141–147]. Pour mieux comprendre le processus et les facteurs de prise de décision en milieu clinique, nous avons assisté à une dizaine de consultations de patients atteints de SEP aux CHU de Rennes et de Nantes et à des réunions entre médecins spécialistes de la SEP pour prendre des décisions sur l'état de santé de certains patients. Nous avons également mené des interviews avec 5 neurologues spécialistes de la SEP sur une période de 6 mois. Une grande partie de ces interviews a été réalisée en distanciel (lors de visio-réunions ou d'appels téléphoniques) en raison de la crise de santé Covid'19.

Cette analyse de l'activité a permis d'identifier des catégories distinctes pour chacun des 3 principaux indicateurs définissant le contrôle de la maladie, à savoir l'EDSS, les nouvelles poussées et les nouvelles lésions à l'IRM (voir Tableau IV.1).

TABLEAU IV.1 – Principaux indicateurs du contrôle thérapeutique de la SEP

| Indicateur   | Catégories d'aggravation  |
|--|---|
| EDSS : progression du niveau de l'incapacité physique du patient | <ul style="list-style-type: none"> <li>• Progression d'un point pour un EDSS entre [1-5.5]</li> <li>• Progression de 0.5 point pour un EDSS <math>\geq</math> 6</li> <li>• Progression de 1.5 point pour un EDSS = 0</li> </ul> |
| Nouvelles poussées : survenue de poussées cliniques              | <ul style="list-style-type: none"> <li>• Aggravation modérée : survenue d'une poussée</li> <li>• Aggravation sévère : survenue de plus d'une poussée</li> </ul>   |
| Nouvelles lésions T2 : apparition de nouvelles lésions           | <ul style="list-style-type: none"> <li>• Aggravation modérée : présence d'une ou deux nouvelles lésions</li> <li>• Aggravation sévère : présence de plus de deux nouvelles lésions</li> </ul>                                   |

L'évolution de ces indicateurs dépend de différents facteurs relatifs au patient, notamment :

- Le sexe

- L'âge au début de la SEP
- La durée de la maladie
- La forme de SEP
- L'EDSS (score permettant de mesurer l'incapacité physique) lors de la dernière visite
- Le nombre de poussées (apparition de nouveaux symptômes ou l'aggravation de symptômes déjà existants) au cours des 12 derniers mois
- Le nombre de lésions T2 sur l'IRM de base
- Les lésions renforcées par le gadolinium (lésions gado) sur l'IRM de base
- Le nombre de nouvelles lésions T2 au cours des 12 derniers mois

Par exemple, un nombre de lésions sur l'IRM supérieur à 10 sur une durée de maladie de moins de 10 ans contribue différemment par rapport au même nombre de lésions sur une durée de la maladie de plus de 10 ans. Nous avons essayé de catégoriser ces facteurs en fonction de la pratique médicale (voir Tableau IV.2) afin de contextualiser les patients atteints de SEP et d'aider à la prise de décisions individualisées plutôt que généralisées.

TABLEAU IV.2 – Catégories des facteurs de la SEP pour la contextualisation des patients

| <b>Facteurs</b>          | <b>Catégories</b>  | <b>Source</b>                                   |
|--------------------------|--|---|
| Sexe                     | <ul style="list-style-type: none"> <li>● Femme (enceinte?)</li> <li>● Homme</li> </ul>   | Catégories communes de SEP (Voir Chapitre III : |
| Forme de SEP             | <ul style="list-style-type: none"> <li>● Syndromes cliniques isolés (SCI)</li> <li>● Récurrente-rémittente (SEP-RR)</li> <li>● Secondairement progressive (SEP-SP)</li> <li>● Primaire progressive (SEP-PP)</li> </ul> | Sclérose en plaques)                            |
| Localisation des lésions | <ul style="list-style-type: none"> <li>● Encéphalique</li> <li>● Tronc cérébral</li> <li>● Moelle épinière</li> </ul>  |   |

TABLEAU IV.2 - Catégories des facteurs de la SEP pour la contextualisation des patients  
(Suite)

| <b>Facteur</b>                                  | <b>Catégories</b>   | <b>Source</b>   |
|---|---|---|
| Traitement suivi                                | <ul style="list-style-type: none"> <li>• Aucun traitement</li> <li>• Première ligne</li> <li>• Deuxième ligne</li> <li>• Troisième ligne</li> </ul> |   |
| Lésions gado                                    | <ul style="list-style-type: none"> <li>• Positive</li> <li>• Négative</li> </ul>  |   |
| EDSS  | <ul style="list-style-type: none"> <li>• [0-2.5]</li> <li>• [3-5.5]</li> <li>• <math>\geq 6</math></li> </ul>                                       | Catégories définies avec des médecins spécialistes de la SEP des CHU de Rennes et de Nantes |
| Nombre de lésions cumulées                      | <ul style="list-style-type: none"> <li>• <math>&lt; 10</math></li> <li>• <math>\geq 10</math></li> </ul>  |   |
| Poussées  | <ul style="list-style-type: none"> <li>• 0</li> <li>• 1</li> <li>• <math>\geq 2</math></li> </ul>   |   |
| Nouvelles lésions au cours des 12 derniers mois | <ul style="list-style-type: none"> <li>• 0</li> <li>• [1-2]</li> <li>• <math>\geq 3</math></li> </ul>   |   |
| Poussées au cours des 12 derniers mois          | <ul style="list-style-type: none"> <li>• 0</li> <li>• 1</li> <li>• <math>\geq 2</math></li> </ul>   |   |

#### IV.4.2 Conception

Cette section décrit les méthodes et les types de visualisation utilisés pour concevoir notre outil d'aide à la décision.



#### IV.4.2.1 Contextualisation : POI / POR

Dans la perspective d'une médecine de précision, il est nécessaire de prendre en compte les facteurs contextuels du patient, c'est-à-dire les facteurs individuels pertinents pour sa prise en charge, et de procéder à une analyse systématique pour identifier les éléments uniques de chaque situation clinique : on parle alors de contextualisation [148]. Pour aider à la prise de décisions médicales pour un patient donné, que nous appelons POI (*Patient Of Interest*), nous utilisons des données provenant d'un ensemble de patients de référence appelés PORs (*Patients Of Reference*). Les PORs représentent les données d'un sous-groupe pertinent de patients partageant des catégories de facteurs de la maladie communes avec le POI. Cette distinction entre les POI et les PORs constitue l'architecture des données dans les applications médicales informatiques. Elle reflète les différentes méthodes de traitement des données liées à la prise de décision médicale pour les sujets de modélisation statistique, les POIs, ainsi que pour la modélisation mathématique liée aux données des PORs. La contextualisation d'un POI consiste à sélectionner ses PORs dans la base de données et à voir son évolution par rapport à ceux-ci. Nous avons utilisé des facteurs validés par des neurologues experts et documentés dans des études antérieures tels que définis dans notre analyse d'activité [125, 143–146] (voir Tableau IV.2) pour la sélection des données des PORs. Comme l'illustre la Figure IV.1, la sélection des PORs d'un POI dépend de différents critères et chaque critère est considéré en fonction de facteurs temporels (âge au début de la SEP, durée de la maladie, etc.).

#### IV.4.2.2 Courbe d'évolution

Pour voir l'évolution de la maladie chez un POI par rapport à ses PORs (autrement dit, pour situer le POI dans le contexte de ses PORs), nous nous sommes inspirés de la conception du CDSS MS Bioscreen [137] quant à l'utilisation des courbes percentiles. Les courbes percentiles fournissent une image complète de la relation entre deux variables, la variable de réponse, pour laquelle nous essayons de trouver les courbes percentiles (ex. EDSS) et la variable explicative (ex. durée de la maladie). Il existe des méthodes paramétriques et non paramétriques pour calculer les percentiles [149], chacune ayant ses propres avantages et inconvénients.

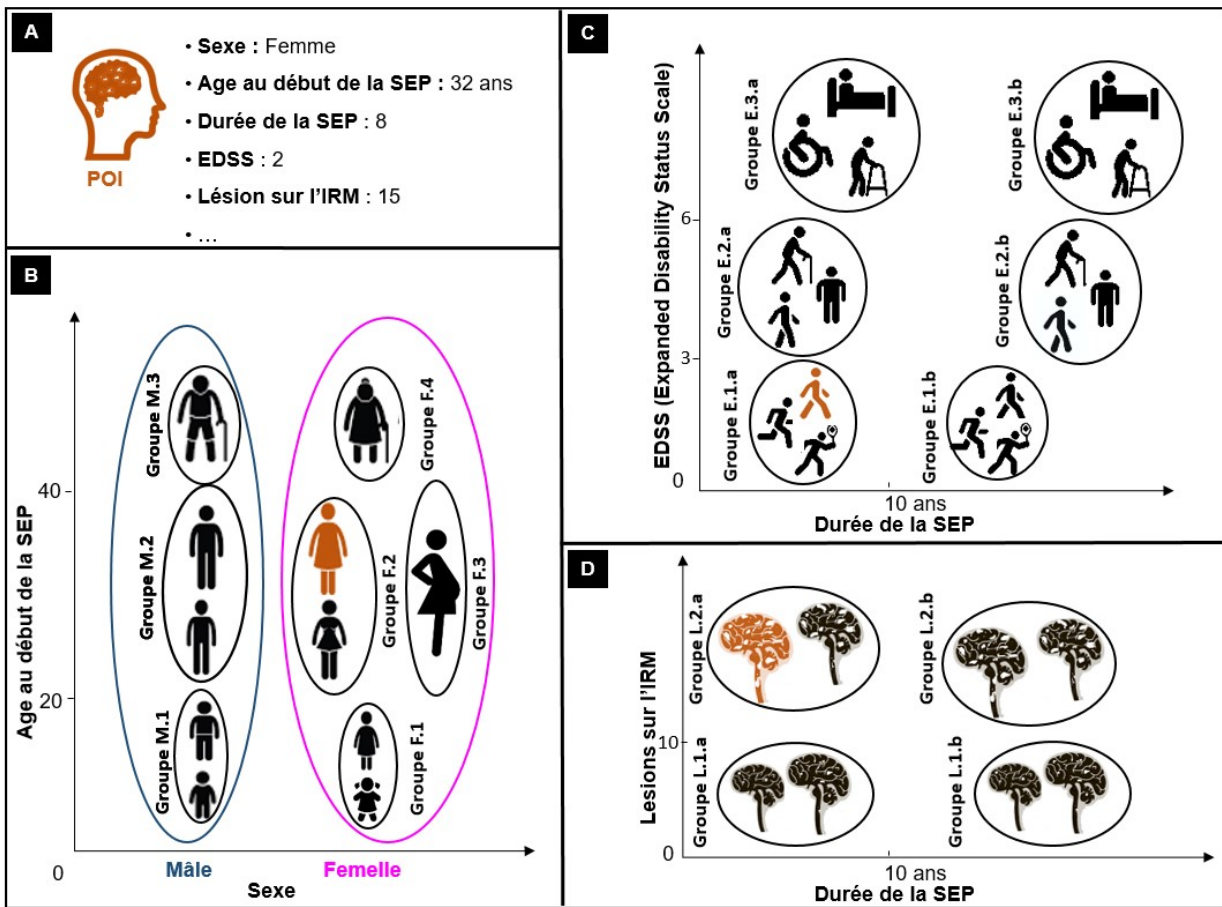


FIGURE IV.1 – Exemple de contextualisation d'un patient atteint de SEP

La contextualisation du POI présenté dans (A) dépend de ses facteurs principaux reconnus de la SEP (par exemple : EDSS, durée de la SEP, nombre de lésions, âge au début de la SEP, etc.). Nous filtrons les données des patients en fonction de l'âge au début de la SEP par rapport au groupe de sexe et si la personne est enceinte. Comme montré dans (B), notre POI est dans le groupe F.2 car c'est une femme dont l'âge au début de la SEP est compris entre [20-40]. Ainsi, nous ne retenons que les données des patients de ce groupe et les filtrons en fonction de l'EDSS, de la durée de la SEP (voir C), du nombre de lésions (voir D), et d'autres facteurs (voir Tableau IV.2)

La méthode feuilles de quantiles (*Quantile Sheets*) [150] est une méthode paramétrique permettant d'estimer les quantiles simultanément. Pour ajuster les percentiles, la méthode des feuilles de quantiles introduit deux paramètres de lissage, l'un dans le sens de la variable explicative et l'autre dans le sens de la probabilité de la variable de réponse. Ces paramètres doivent être définis par l'utilisateur par inspection ou par d'autres méthodes. Les modèles additifs généralisés de localisation, d'échelle et de forme (GAMLSS - *Generalized Additive Model for Location, Scale and Shape*) [151] sont des modèles de régres-

sion semi-paramétrique dont la distribution de la variable réponse peut être modélisée en proposant plusieurs méthodes pour lisser les courbes percentiles.

Pour automatiser la sélection du modèle le plus approprié pour la construction des courbes de croissance, la méthode LMS (*Lambda-Mu-Sigma*) a été développée. Cette méthode est capable de déterminer la distribution de la variable de réponse avec tous les paramètres nécessaires (degrés de liberté, paramètre de puissance pour la transformation, etc.) pour l'ajuster à une variable explicative afin de créer des courbes percentiles. La Régression Quantile (RQ) [152] est une méthode non paramétrique, dans laquelle chaque centile est estimé séparément. Le modèle RQ est flexible, aucune distribution n'est supposée pour la variable de réponse, ce qui réduit les biais produit par une distribution erronée.

Notre objectif principal est de fournir les bonnes informations et le calcul à la demande tout en garantissant l'ergonomie de l'utilisation. Ainsi, nous avons éliminé les méthodes des feuilles de quantiles, car les paramètres doivent être choisis par l'utilisateur à travers l'analyse des courbes obtenues. Nous avons également abandonné la méthode LMS car elle prend beaucoup de temps à calculer et ne permet pas de réduire automatiquement les biais qui peuvent être causés par l'hypothèse d'une distribution erronée. Nous avons donc opté pour la méthode RQ afin de voir comment se situe l'EDSS d'un patient d'intérêt par rapport à ses PORs. Comme il s'agit d'une méthode non paramétrique, le médecin n'aura pas à changer les paramètres. De plus, elle ne requiert pas un long temps de calcul, permettant ainsi un calcul à la demande et en temps réel.

### IV.4.2.3 Projection

La projection des données est une méthode permettant de faire des prévisions sur l'évolution de la maladie chez le POI au cours des années à venir. Le processus de projection des données (voir Figure IV.2) commence par le filtrage appliqué pour contextualiser le POI. Les PORs sélectionnés seront ensuite considérés du temps  $t_0$  (le moment où ils représentaient des facteurs de la maladie, similaires à ceux du POI pour les critères de contextualisation) à un temps  $t_i$  (par exemple, un an plus tard). L'objectif est de montrer sous tel traitement la probabilité raisonnable de demeurer dans un état donné à  $t_i$  pour les 3 indicateurs principaux (voir Tableau IV.1) qui définissent le contrôle de la maladie.

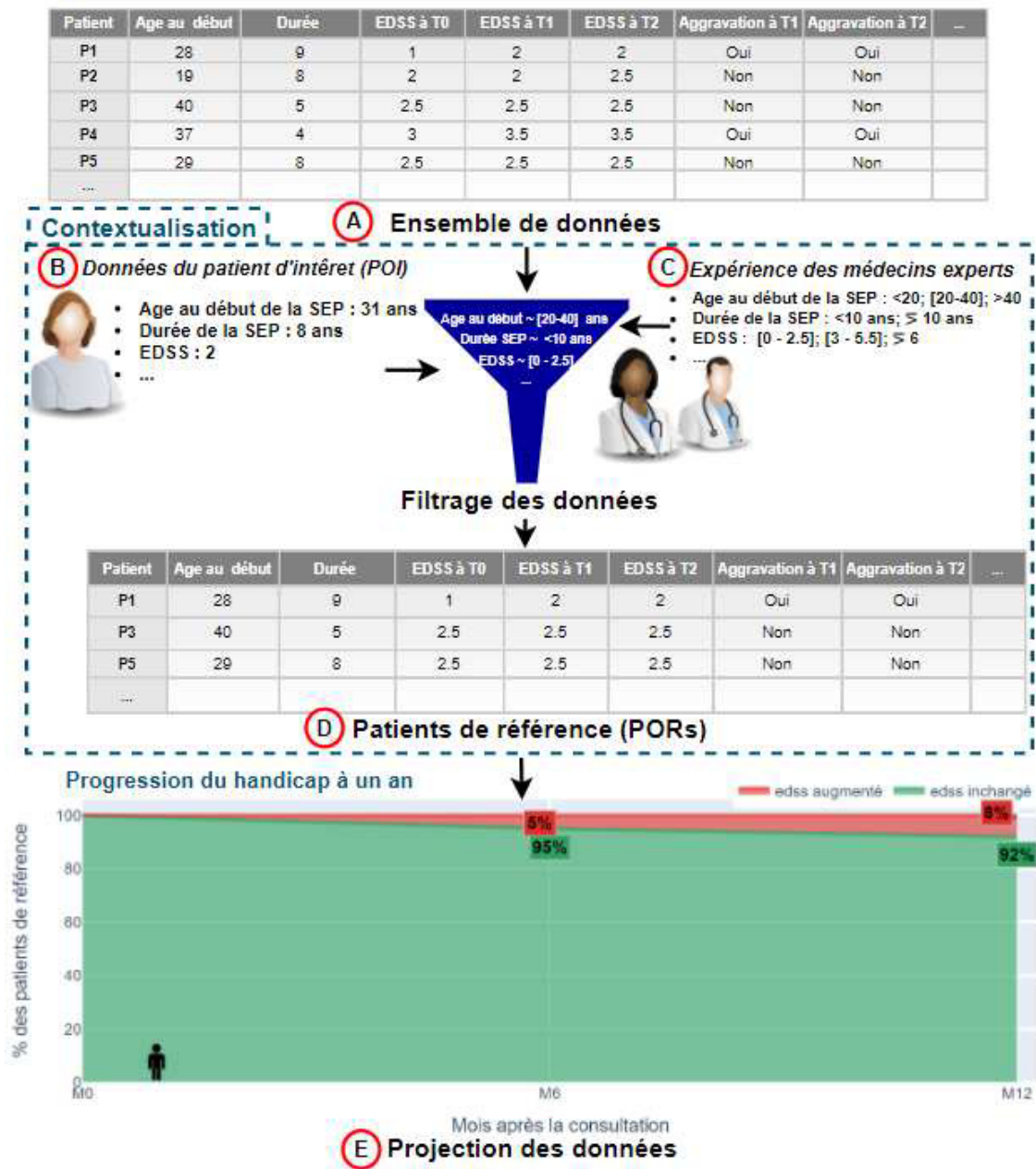


FIGURE IV.2 – Processus de projection de données

Le processus de projection de données repose sur le filtrage de l'ensemble de données (A), en fonction des données du POI (B), et de l'expérience des médecins experts (C). Les données filtrées (PORs) dans (D) sont ensuite utilisées pour visualiser l'évolution potentielle de la maladie du POI en fonction de la projection des données, exemple de la progression du handicap dans (E).

Nous nous sommes appuyés sur les travaux de J. Flach [10, 64] sur le développement d'interfaces écologiques pour orienter la conception des graphiques de visualisation des projections de données. Ces travaux nous ont guidés dans la création des représentations visuelles qui englobent les interactions complexes entre les données médicales et leur

contexte, tout en étant intuitives pour les utilisateurs. Nous avons utilisé des couleurs qui reflètent intuitivement les différents niveaux d'aggravation des principaux indicateurs de l'évolution de la maladie. Par ailleurs, l'épaisseur des couleurs dans les graphiques donne une indication visuelle de l'importance respective de chaque niveau d'aggravation.

### **IV.4.3 Prototypage**

Cette section décrit les différentes évaluations réalisées pour développer notre CDSS et présente les différentes fonctionnalités du prototype fonctionnel final.

#### **IV.4.3.1 Évaluation en vue d'un développement basé sur l'utilisateur**

Pour faire des prévisions sur l'évolution et l'activité de la maladie chez un patient donné, nous avons conçu un modèle illustrant la structure de décision identifiée lors de notre analyse de l'activité et basé sur la conception décrite précédemment. Ce modèle a servi de base à un premier prototype de MS Vista. Ce prototype a évolué au fil du temps grâce à des évaluations menées auprès des médecins spécialistes de la SEP afin de déterminer s'il répondait à leurs attentes. Nous avons procédé à différents types d'évaluations. Une évaluation formelle dans laquelle 4 médecins étaient invités à effectuer des exercices pour faire des prédictions sur l'évolution de la SEP et le traitement à recommander aux patients avec et sans l'utilisation de MS Vista. Pour chaque exercice, les médecins ont été invités à indiquer les facteurs utilisés pour faire les prédictions. À la fin de chaque test, les médecins étaient invités à répondre à un questionnaire afin d'évaluer leur satisfaction et l'utilité du prototype proposé pour la prise de décision médicale. En guise de conclusion des résultats des premières activités de cette évaluation, MS Vista peut agir comme un moyen d'aide à la décision et d'alerte pour éviter d'induire en erreur certains neurologues en raison de la multiplicité des facteurs liés à cette maladie. Les résultats du questionnaire (voir Annexe B : Résultats du questionnaire d'évaluation de MS Vista pour le développement centré sur l'utilisateur) ont montré que  $\frac{3}{4}$  des participants ont trouvé l'application ergonomique. Tous les participants étaient satisfaits de l'application. En ce qui concerne l'exhaustivité des critères de sélection du PORs, les réponses étaient disparates, deux médecins les trouvaient insuffisants, un médecin avait une position neutre alors qu'un participant les jugeait largement suffisants. Quelques suggestions de modifi-

cations ont été proposées, telles que l'âge au début de la maladie et la durée de la maladie au lieu de la date de naissance du patient et de la date de début de la maladie, la mesure de l'aggravation de l'EDSS, etc. La possibilité d'inclure d'autres facteurs et d'autres types de traitement a été discutée. Ainsi, nous avons réalisé une autre itération du processus de conception pour mettre en place les suggestions des participants. Des évaluations informelles lors de réunions ou d'appels téléphoniques avec des médecins ont eu lieu afin de tester MS Vista et de s'assurer de certains éléments par exemple par rapport à leurs besoins dans la phase de communication avec le patient. Tout cela a conduit au développement du prototype fonctionnel du CDSS MS Vista décrit dans la section suivante.

#### IV.4.3.2 Fonctionnalités de MS Vista

Le CDSS MS Vista offre deux options d'utilisation. La première concerne l'aide à la décision pour les nouveaux patients (voir figure IV.3).

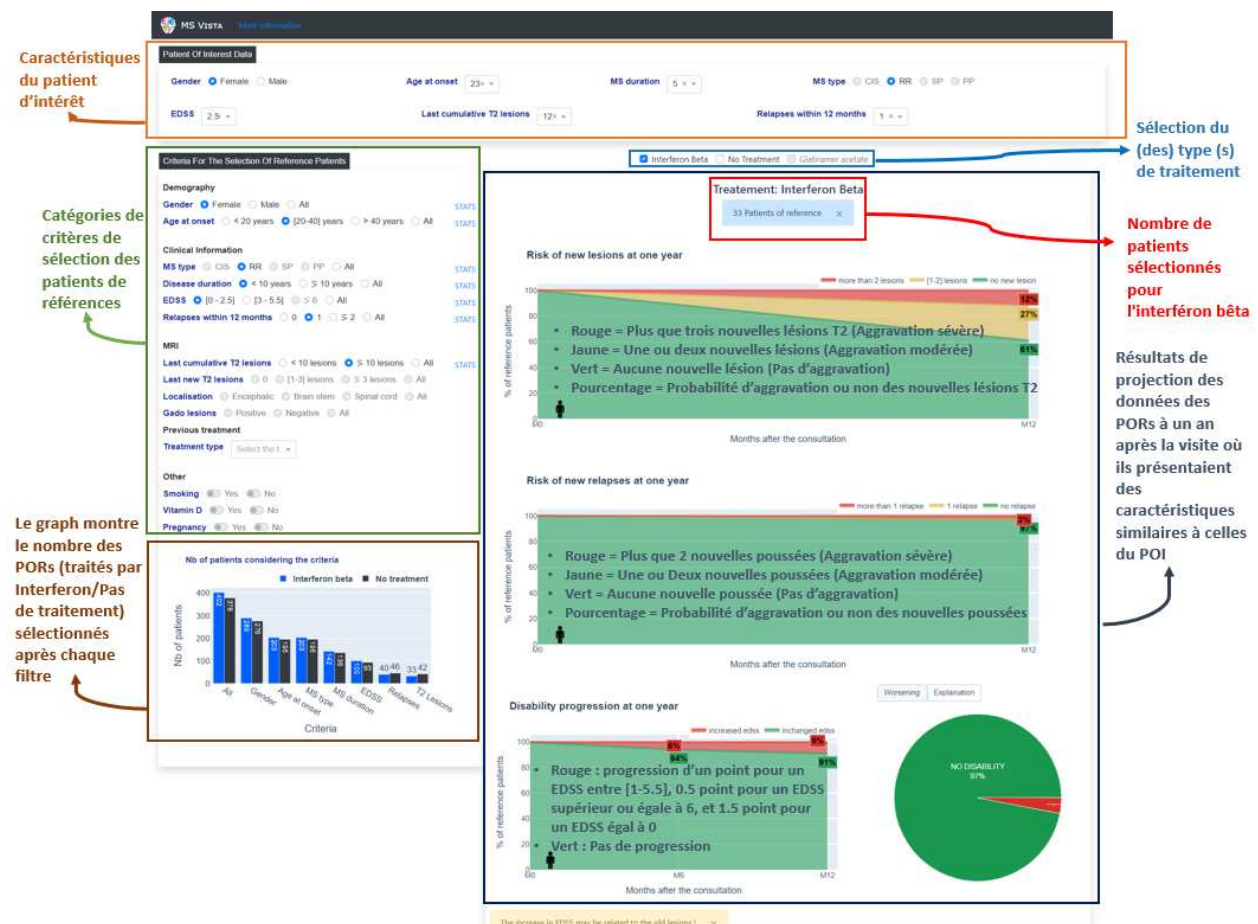


FIGURE IV.3 – MS Vista, interface de contextualisation et de projection des données



La seconde concerne les patients dont les données figurent déjà dans la base de données. En plus des fonctionnalités de la première option, la seconde option permet une visualisation graphique de l'historique des changements des lésions T2 et de la trajectoire de l'EDSS du POI. En utilisant les données enregistrées du POI, MS Vista illustre la trajectoire de son EDSS dans le contexte d'une distribution percentile basée sur l'EDSS de ses PORs (voir Figure IV.4).

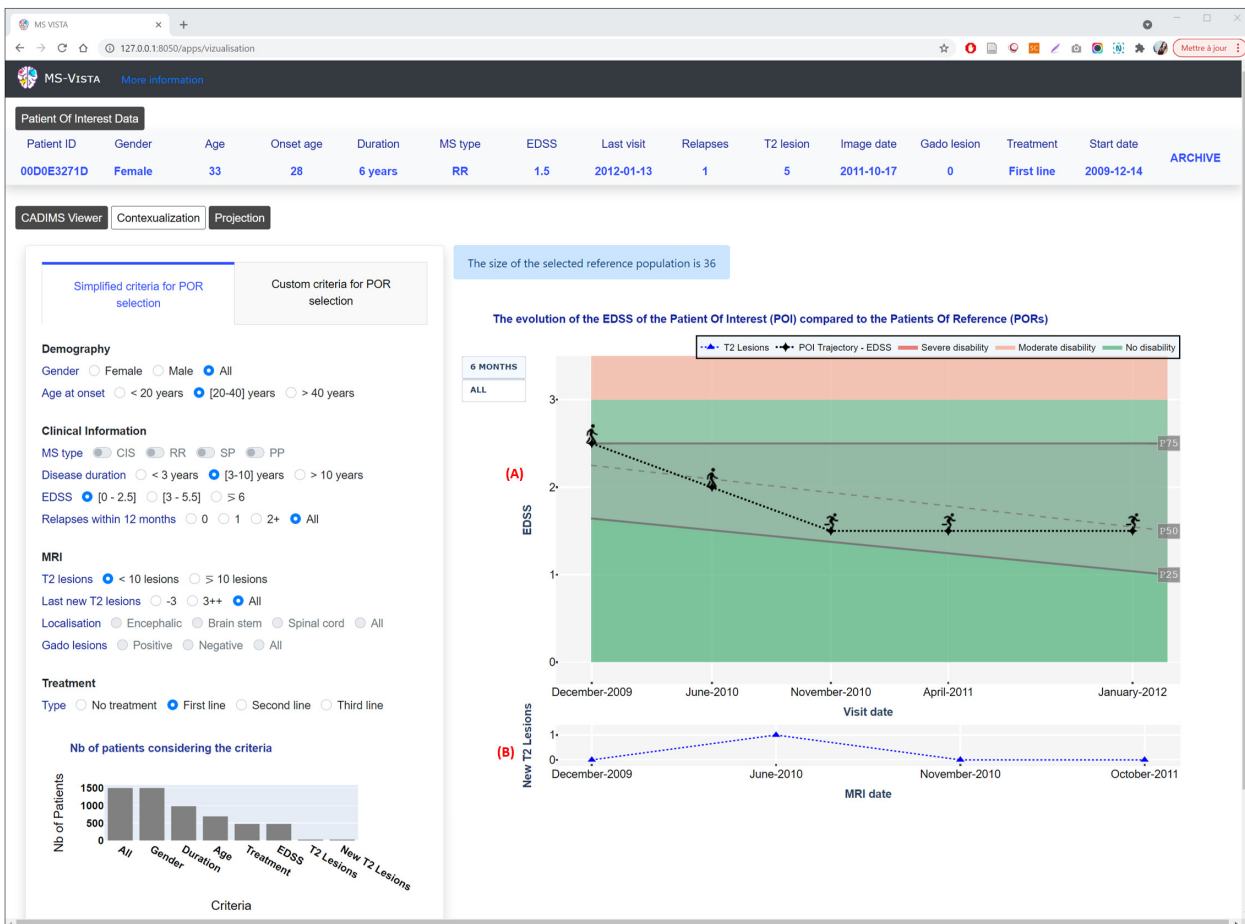


FIGURE IV.4 – Visualisation de l'évolution de l'EDSS d'un POI par rapport à ses PORs (au fil du temps) dans MS Vista (source Ed-Driouch et al. [2])

(A) La trajectoire de l'EDSS du POI dans le contexte d'une distribution de percentile des PORs au cours de la durée de la SEP. Le POI est présenté par une icône d'humain qui change en fonction de la valeur EDSS (voir Figure III.3 du Chapitre 3 : Sclérose en plaques) pour être plus illustratif pour les patients. Deux vues sont disponibles, l'EDSS une fois tous les 6 mois pour plus de clarté, et tous les EDSS enregistrés pour le patient pour une analyse plus approfondie. Une courbe des nouvelles lésions T2 à chaque IRM du POI est aussi affichée (B) pour aider à déterminer si le POI évolue normalement.

Cette évaluation normative des trajectoires potentielles du patient pourra aider à obtenir une estimation personnalisée du contrôle de l'activité de la maladie.

Pour guider les médecins non-experts et faire gagner du temps aux experts, MS Vista offre une sélection automatique des PORs (voir Figure IV.5) par des filtres prédéfinis de catégories de sélection des PORs. Ceux-ci sont basés sur des mesures prédictives de l'activité clinique et IRM de la maladie selon les neurologues experts et la littérature, comme indiqué dans notre analyse de l'activité. Une sélection personnalisée (voir Figure IV.6) est aussi possible pour permettre aux médecins de calibrer les catégories de sélection des PORs (par exemple pour ne sélectionner que des PORs dont l'EDSS est égal à 2 au lieu de toute la catégorie [0-2.5]).

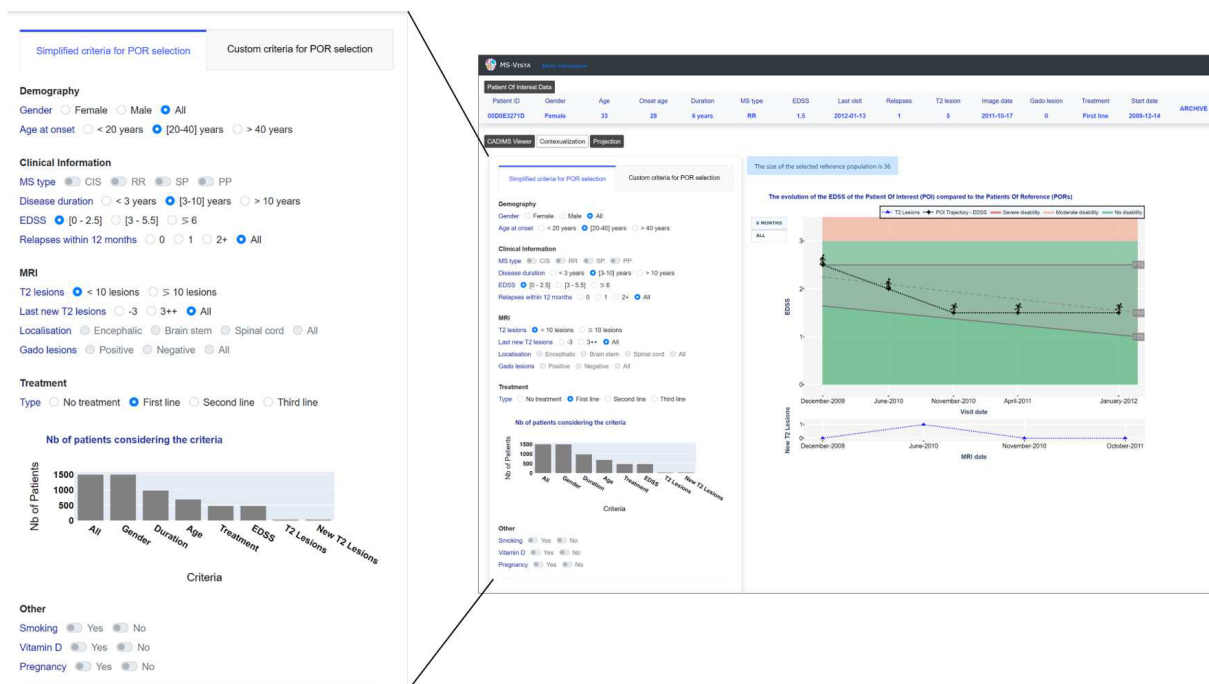


FIGURE IV.5 – Sélection automatique des PORs dans MS Vista par des filtres prédéfinis de catégories de sélection des PORs

MS Vista permet au praticien de voir la qualité de la base de données dans son ensemble et la qualité des données des PORs sélectionnées pour chaque POI (Voir Figure IV.7). Il montre les statistiques descriptives des données par rapport à chaque facteur utilisé pour la sélection des PORs sous forme de diagramme en boîte (boite à moustaches). Les diagrammes en boîtes affichent les valeurs médiane, moyenne, min, max et les quartiles pour visualiser la distribution des données numériques et l'asymétrie. Cela permet-



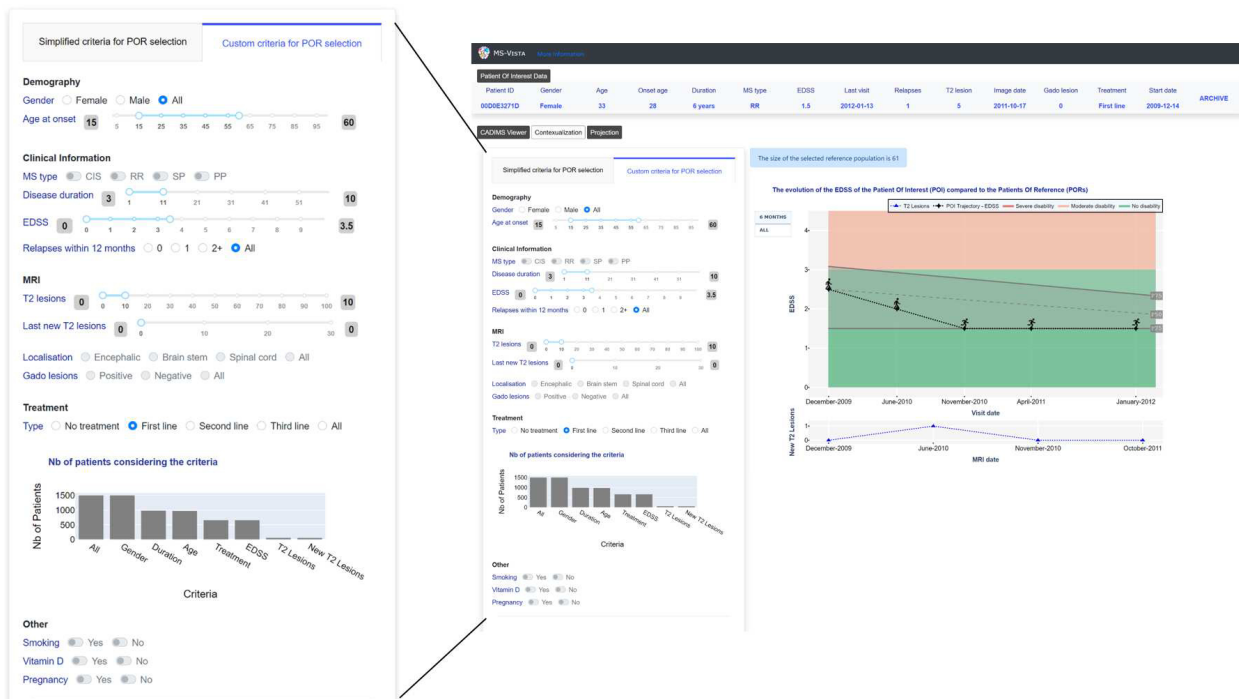


FIGURE IV.6 – Sélection personnalisée des PORs dans MS Vista permettant de calibrer les catégories de sélection des PORs

tra aux médecins de voir si les données de chaque facteur représentent la catégorie de contextualisation correspondante. Ainsi, le praticien peut examiner la qualité des données des PORs de chaque POI et décider si elles sont conformes et appropriées pour la contextualisation du POI avant de procéder à tout traitement. Par exemple, un patient qui a développé 11 lésions devrait être contextualisé dans la catégorie « >10 lésions » et le graphique montre que la majorité des patients de cette catégorie sont affichés comme ayant développé plus de 30 lésions. Cela peut amener le médecin à ne pas prendre en compte les résultats fournis par le CDSS dans ce cas, car les données ne présentent pas vraiment le contexte du patient. MS Vista fournit des prévisions sur l'évolution de la SEP chez un POI en fonction de l'aggravation des facteurs prédictifs de l'évolution de la maladie (voir Tableau IV.1) chez ses PORs. Comme le montre les Figures IV.8 et IV.9, MS Vista illustre le risque potentiel de l'aggravation de ces facteurs chez le POI à partir du pourcentage de leurs niveaux d'aggravation chez les PORs au cours de l'année suivant la visite où ces derniers représentaient des facteurs de maladie similaires au POI (projection des données des PORs sur un an). De plus, MS Vista présente les projections de données par rapport à chaque option de traitement (voir Figure IV.10). Il s'agit d'envisager différents scénarios

thérapeutiques illustrant les avantages potentiels des différentes options thérapeutiques. Ainsi, enrichir la communication entre le médecin et le patient sur le choix du traitement.

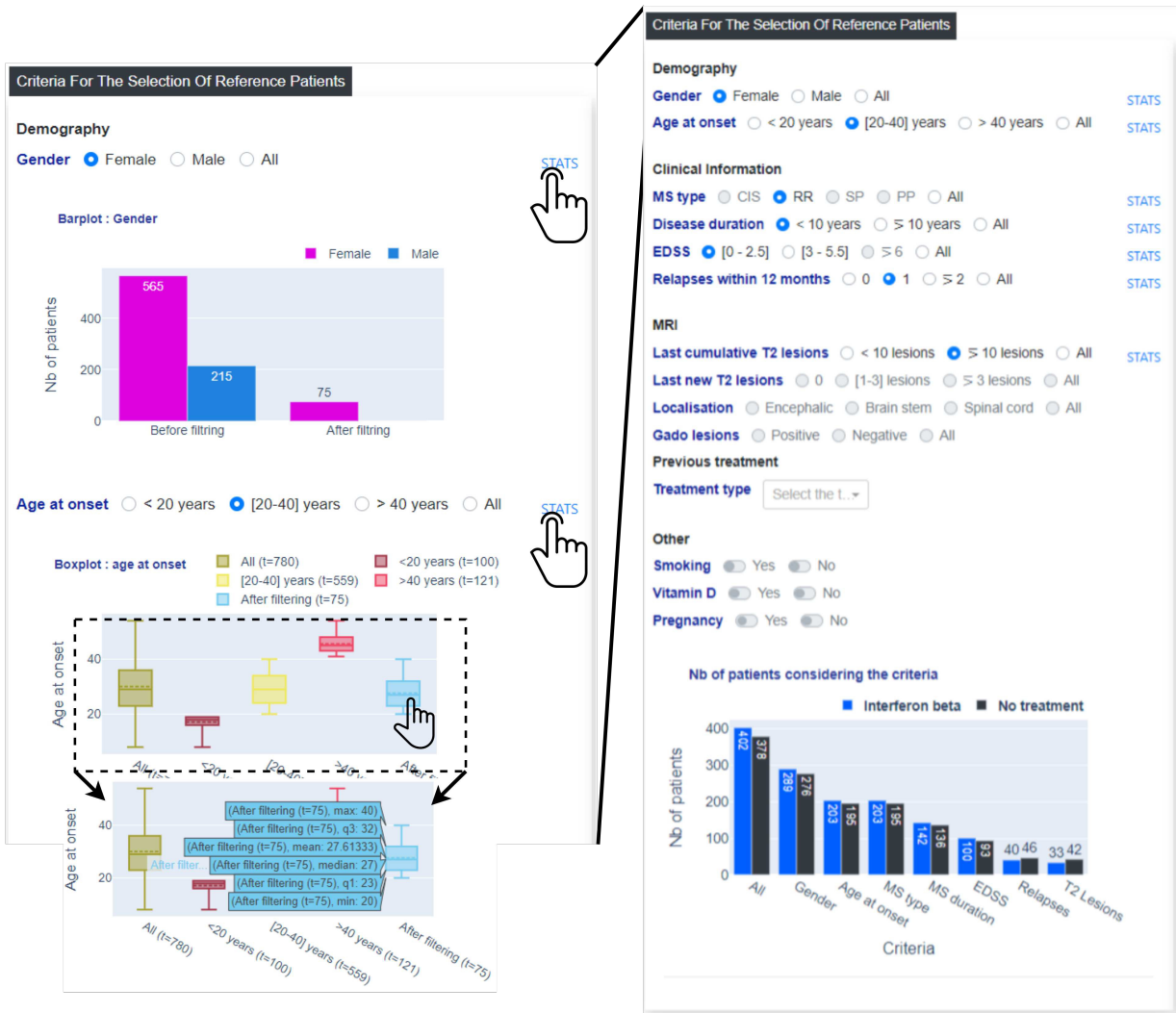


FIGURE IV.7 – Statistiques descriptives fournies par MS Vista sur les critères de sélection des PORs

En cliquant sur le bouton Stats associé à chaque facteur, MS Vista affiche les statistiques descriptives du facteur au moyen de diagrammes à barres ou de boîtes à moustaches interactifs. Le passage du curseur sur ces dernières affiche automatiquement les descriptions statistiques (Min, Max, Médiane, etc.).

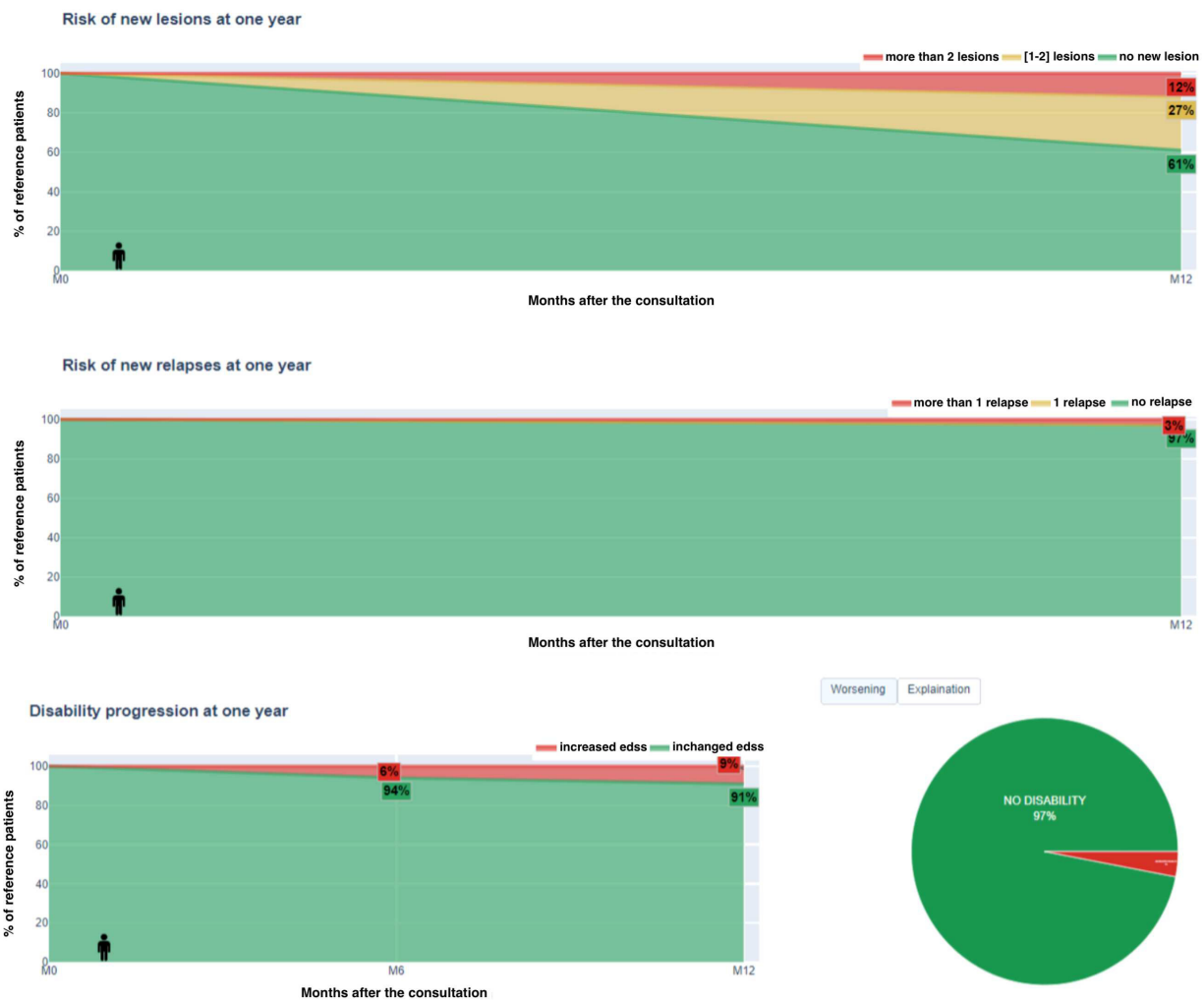


FIGURE IV.8 – Prévion de l'évolution potentielle de la SEP via les projections de données dans MS Vista

*La projection des données des PORs dans MS Vista permet d'obtenir des graphiques illustrant l'activité potentielle de la SEP. Cela se traduit par la mesure du pourcentage de nouvelles poussées (0 / 1 / >1), l'accumulation de nouvelles lésions T2 sur l'IRM (0 / [1,2] / >2) et l'aggravation de l'EDSS (aggravation / pas d'aggravation).*

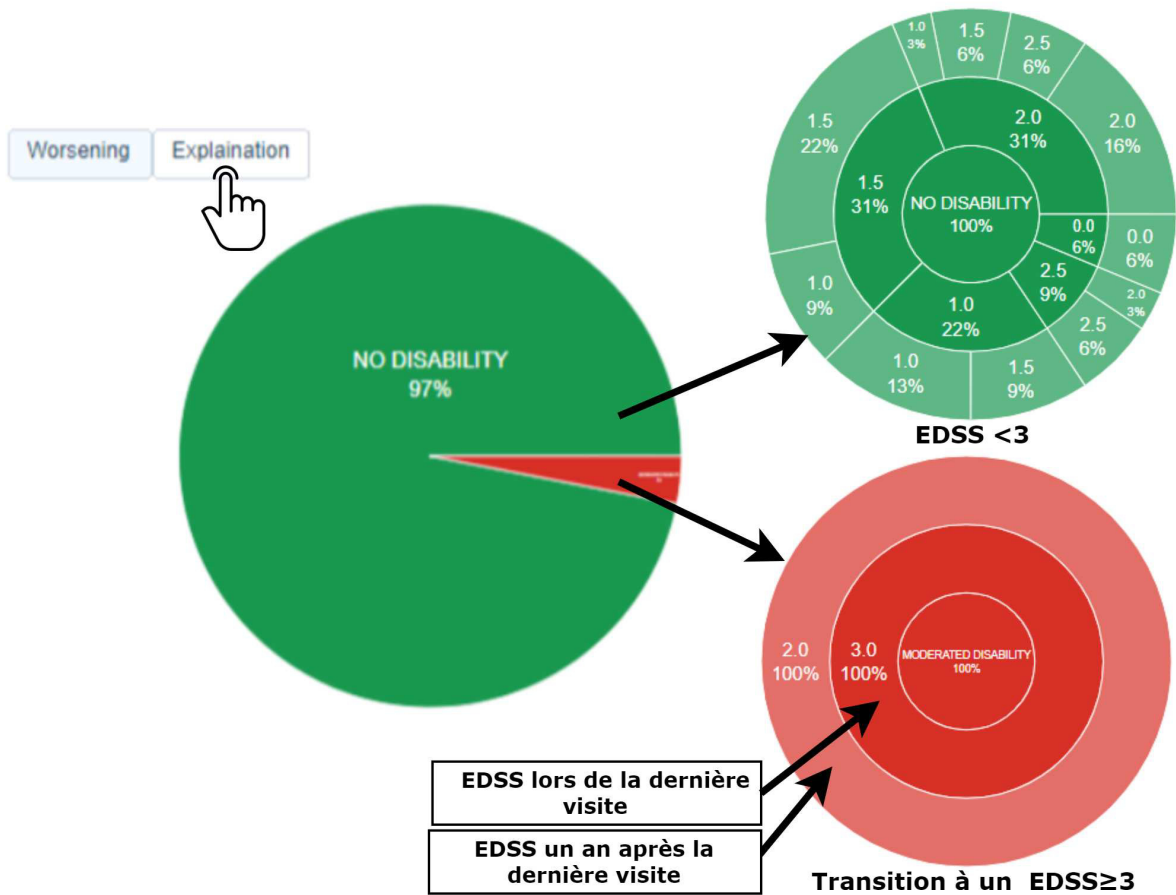


FIGURE IV.9 – Mode dual de visualisation de l'aggravation de l'EDSS dans MS Vista

*Le diagramme circulaire fourni par MS Vista offre un mode dual de visualisation de l'aggravation de l'EDSS. Le premier niveau de visualisation montre le pourcentage d'aggravation du handicap (passage d'un EDSS < 3 à un EDSS ≥ 3 [125]) un an plus tard. Le deuxième niveau affiche le nombre de patients correspondant à chaque EDSS, avec plus de détails pour fournir une représentation plus précise.*

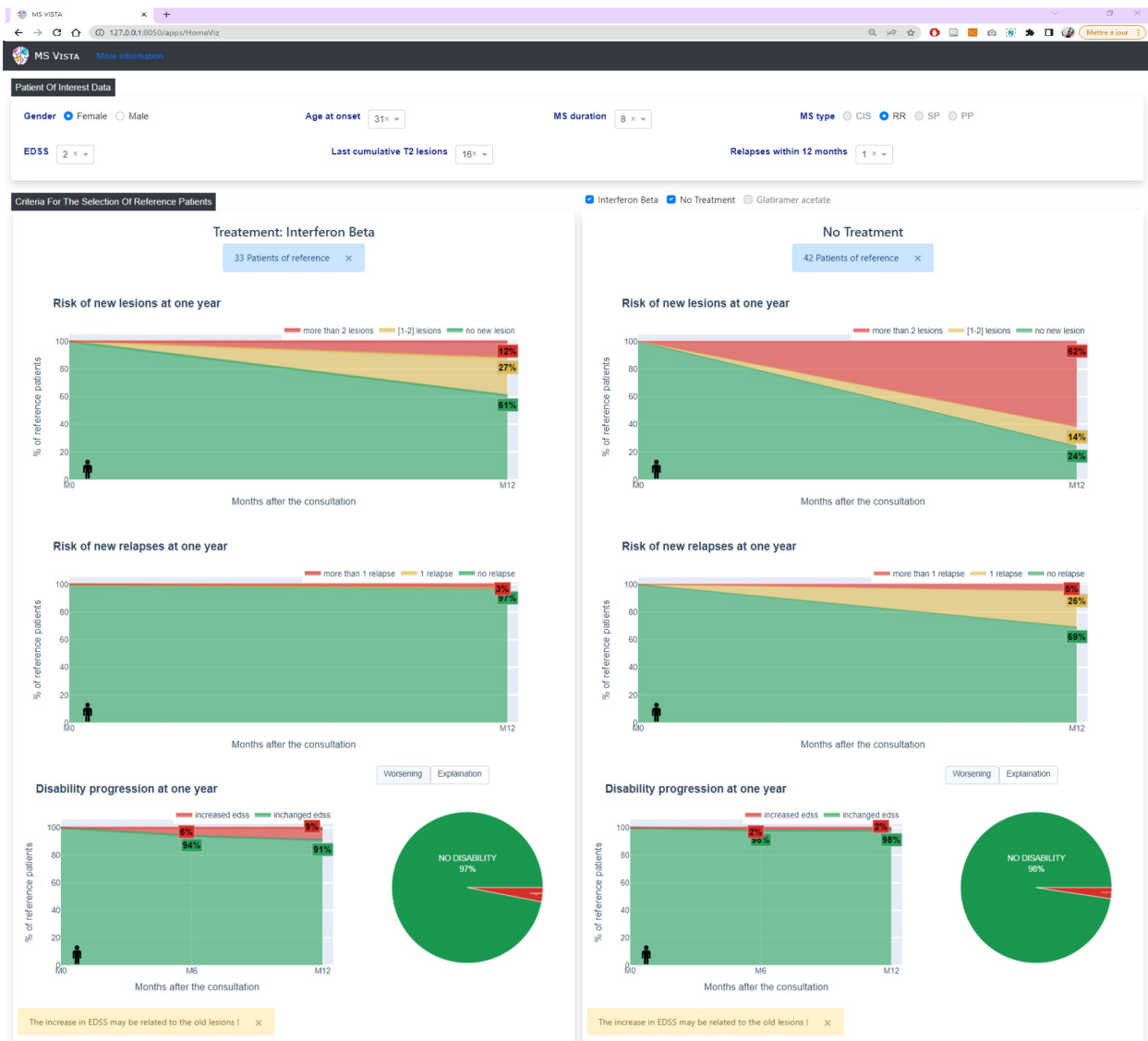


FIGURE IV.10 – Illustration comparative des effets potentiels de chaque option thérapeutique dans MS Vista (source Ed-Driouch et al. [2])

*MS Vista fournit une illustration comparative de l'efficacité potentielle du traitement Peginterferon bêta-1a par rapport à l'absence du traitement pour chaque POI, en tenant compte des facteurs de sa maladie.*

Pour accompagner les processus de réflexion et de communication [59], MS Vista propose des composants interactifs et des graphiques pour la visualisation à différents niveaux. D'une part, une représentation simplifiée conçue pour les échanges avec les non-experts, notamment les patients, et d'autre part, un niveau plus expert pour aider les médecins dans leur prise de décision. Cette approche est basée sur les besoins des mé-

decins dans les phases de prise de décision et de communication avec les patients, et s'appuie sur les contributions de J. Flach dans le domaine de la visualisation de données [10]. Plus précisément, l'introduction d'un deuxième niveau de visualisation pour améliorer la compréhension des visuels d'aide à la prise de décision relative aux maladies cardiovasculaires. Au cœur de notre approche, nous avons tiré parti des travaux de Flach pour concevoir deux niveaux de visualisation. Un premier niveau est destiné au médecin pour lui fournir des informations sur les données utilisées, des statistiques sur les données des PORs (voir Figure IV.7) et sur leur conformité avec les données du POI, et les mesures qui sous-tendent les affichages (voir l'exemple du graphique circulaire sur l'aggravation de l'EDSS dans la Figure IV.9). Un deuxième niveau est conçu pour concentrer la communication entre le médecin et le patient sur les effets potentiels de chaque option de traitement (voir Figure IV.10).

#### **IV.4.4 Évaluation de MS Vista comme outil support de la communication médecin patient**

Le développement de MS Vista s'est appuyée sur l'expérience des médecins et sur leurs besoins dans la phase de communication avec le patient, afin de fournir une solution axée sur les données qui aidera à la prise de décision médicale dans la SEP et facilitera la communication médecin-patient. Pour tester la qualité de cette solution et valider ses avantages, il est nécessaire d'évaluer l'expérience des utilisateurs cibles (médecins et patients). À cette fin, nous avons élaboré un protocole d'évaluation (n° IRB : IORG0011023) sur la base de la méthode de C. Lallemand et G. Gronier [153]. Cette section décrit les différents volets du protocole d'évaluation. Il est à noter qu'en raison de certaines contraintes, l'évaluation a été limitée à un groupe restreint de participants, à savoir 3 patients et 2 médecins (3 consultations).

##### **Hypothèse**

MS Vista aide à la communication médecin-patient lors des visites de suivi de la SEP, parce qu'elle représente un support de discussion pour le médecin pour l'aider à expliquer son pronostic

### Type du test utilisateur

Test en contexte naturel, évaluation du système dans l'endroit où l'utilisateur effectue ses interactions avec le système.

- CHU de Rennes, Clinique de la sclérose en plaques.
- CHU de Nantes, Hôpital nord Laennec.

### Participants

- Médecins de sclérose en plaques
- Patients atteints de sclérose en plaques récurrente-rémittente (SEP-RR) sans traitement ou sous interféron bêta (ou pour lesquels ce traitement pourrait être prescrit).

### Durée du test

Environ 40 minutes par consultation.

TABLEAU IV.3 – Précisions sur la durée de l'évaluation

| Étape de la passation                | Durée  | Durée cumulée |
|--------------------------------------|--------|---------------|
| Rencontre du médecin et instructions | 5 min  | 5 min         |
| Accueil du patient et instructions   | 10 min | 15 min        |
| Réalisation des scénarios d'usage    | 20 min | 35 min        |
| Passation du post-questionnaire      | 5 min  | 40 min        |

### Équipements

- Ordinateur
- Caméra
- Documents papiers (Pré-questionnaire générique/ formulaire de consentement, Post-questionnaire) (Voir en Annexe B : Évaluation de MS Vista comme outil support de la communication médecin-patient)

### **Mesures**

- Niveau d'interaction médecin-patient avec et sans utilisation du système : nombre et type de questions posées et répétées par le patient, difficultés rencontrées par le médecin en expliquant les résultats obtenus par l'application, etc.
- Mesures subjectives : questionnaires sur l'avis et la satisfaction des patients et des médecins à l'égard de l'application.

### **Conditions expérimentales**

Les médecins sont invités à utiliser l'application avant de recevoir le patient, afin de vérifier l'adéquation des données, la qualité des résultats et la faisabilité de les présenter au patient concerné. Le médecin demande ensuite au patient s'il accepte de participer à l'évaluation. Dans l'affirmative, il signe le formulaire de consentement.

### **Description du protocole**

L'expérience, d'une durée de 40 min environ, est réalisée en contexte naturel (lieu de consultation de suivi), en deux étapes. La première consiste à faire la consultation sans utiliser l'application. La seconde sert à utiliser l'application pour montrer l'évolution d'un ensemble de patients ayant des facteurs cliniques et démographiques similaires au patient concerné par la consultation, afin d'aider le médecin à expliquer le pronostic au patient. À la fin de la consultation, les patients évaluent leur expérience de l'utilisation de l'application comme support de discussion au travers d'un questionnaire à six questions fermées sur l'échelle de Lickert et une question ouverte (Voir en Annexe B : Évaluation de MS Vista comme outil support de la communication médecin-patient, Mesure de l'utilisabilité du système – formulaire patient). Le médecin évalue son expérience avec l'application dans la communication avec le patient au travers d'un questionnaire à cinq questions fermées sur l'échelle de Lickert (Voir en Annexe B : Évaluation de MS Vista comme outil support de la communication médecin-patient, Mesure de l'utilisabilité du système - formulaire médecin (2)), et la qualité de l'interaction avec l'application par la version française abrégée du questionnaire AttrakDiff [154] (Voir en Annexe B : Évaluation de MS Vista comme outil support de la communication médecin-patient, Mesure de l'utilisabilité du système - formulaire médecin).



## Évaluation

Les critères d'évaluation qui restreignaient la participation aux seuls patients atteints de SEP-RR qui devaient être sous interféron bêta ou à qui ce traitement pouvait être prescrit, ont limité le nombre de participants potentiels. Le faible nombre de visites de ce groupe de patients a créé des obstacles à la poursuite de cette évaluation. Nous n'avons pu réaliser notre évaluation qu'avec 3 patient(e)s et 2 médecins. Par ailleurs, le contrat d'utilisation des données a expiré durant l'évaluation et son renouvellement a pris du temps. Néanmoins, les retours des patients et des médecins lors des 3 évaluations ont été remarquablement positifs, reflétant un fort enthousiasme de la part des utilisateurs pour l'intégration de l'outil dans les consultations. En outre, ces retours ont suscité des demandes d'intégration de données sur d'autres traitements, afin de permettre des choix thérapeutiques plus précis et plus éclairés.

Le CDSS MS Vista est une preuve de concept qui a aidé à obtenir un financement de 8 271 646 € sur cinq ans de l'Agence Nationale de la Recherche (ANR) dans le cadre de son 3<sup>ème</sup> PIA, intégré dans le plan France 2030 sous la référence [ANR-21-RHUS-0014]. Le projet PRIMUS offre la possibilité d'une présentation plus personnalisée de l'efficacité à long terme des médicaments de fond approuvés, grâce à l'accès aux données d'essais cliniques randomisés et aux cohortes longitudinales complètes de haute qualité de patients atteints de SEP. Environ 3 000 patients de la cohorte OFSEP-HD et des données d'essais cliniques de phase III des laboratoires Biogen (4 200 patients) et Merck (2 400 patients) [155]. Cette démarche facilitera la reprise des évaluations avec plus de participants. Ceci grâce à l'alimentation de MS Vista avec plus de données et de traitement de la SEP.

## Discussion

Ce chapitre décrit le prototype fonctionnel de CDSS pour la SEP, MS Vista, fruit d'une collaboration entre des professionnels de la santé spécialisés dans la SEP, des chercheurs en informatique et facteurs humains. MS Vista repose sur la distinction entre le POI, le patient pour lequel on est amené à prendre une décision médicale, et les PORs, des patients présentant des facteurs pertinents de la maladie similaire à ceux du POI. Ces facteurs sont validés dans la littérature [125, 143–146] et par des neurologues experts. MS Vista est basé

sur les données d'un essai clinique de phase III (NCT00906399) pour contextualiser un POI (recherche de ses PORs). Il permet des projections de données des PORs pour aider les cliniciens à concevoir des prévisions sur l'évolution des principaux facteurs prédictifs de l'activité et de la progression de la SEP : survenue de nouvelles poussées, présence de nouvelles lésions sur l'IRM et l'aggravation de l'EDSS. En outre, il fournit des visualisations des résultats de la projection à un an par rapport à chaque option de traitement : peginterféron bêta-1a versus absence de traitement.

Nous avons conçu MS Vista en tenant compte des limites des CDSS déjà développés pour aider à la prise de décision dans la SEP, présentées dans la section « État de l'art » de ce chapitre. Nous nous sommes appuyés sur les connaissances des médecins experts en SEP pour concevoir notre CDSS. MS Vista permet de contextualiser les patients atteints de SEP afin d'aider à la personnalisation des décisions médicales. Il permet au médecin d'examiner la qualité des données pour déterminer si les indications de décision fournies sont favorables et peuvent être montrées au patient. De plus, il a le potentiel d'enrichir la communication médecin-patient grâce à des visualisations interactives montrant comment différents facteurs contribuent à la progression de la maladie, tout en illustrant les avantages potentiels de chaque option de traitement.

Une des limites du CDSS MS Vista est le manque de certaines informations dans les données des PORs telles que les biomarqueurs prédictifs (chaîne légère des neurofilaments, protéine fibrillaire acide gliale, etc.), ou encore des facteurs tels que l'état émotionnel et cognitif, la qualité de vie, etc. Une autre limite du CDSS actuel est l'intégration d'autres traitements approuvés pour la SEP. Pour cette raison, il est important que les CDSS soient conçus comme des assistants, rendant visible les limites de leur raisonnement, afin que le médecin puisse continuer à intégrer de son côté les autres paramètres dans sa projection.

Nous avons utilisé MS Vista pour illustrer l'efficacité potentielle de l'interféron bêta par rapport à l'absence de traitement. Les résultats ont montré que l'efficacité du traitement est fortement liée aux facteurs de la maladie [2]. En représentant l'effet général de l'interféron bêta par rapport à l'absence de traitement (placebo), nous avons remarqué que 48% des 378 patients sous placebo ont développé plus de 2 lésions contre 14% des 402 patients sous interféron. D'autre part, en prenant uniquement le cas des patients dont l'âge au début de la maladie est inférieur à 20 ans et qui présentent plus de 10 lésions

sur l'IRM, 45% des patients sous interféron bêta ont développé plus de 2 lésions un an plus tard, indiquant ainsi l'absence d'effet du traitement sur cette sous-population. Cela illustre l'importance de faire des choix thérapeutiques individualisés.

# PRÉDICTION HYBRIDE POUR L'AIDE À LA DÉCISION MÉDICALE

## Sommaire du chapitre

|  |    |
|--|----|
| Introduction . . . . .   | 78 |
| V.1 État de l'art . . . . .  | 78 |
| V.2 Hypothèse . . . . .  | 79 |
| V.3 Nouvelle approche de l'intelligence hybride . . . . .                                      | 80 |
| V.3.1 Collaboration médecin-algorithme . . . . .   | 80 |
| V.3.2 Évaluation de l'approche d'intelligence hybride médecin-algorithme . . . . .             | 81 |
| V.3.3 Classes déséquilibrées : pipeline pour la création de modèles de<br>prédiction . . . . . | 82 |
| V.3.3.1 Séparation des données . . . . .   | 83 |
| V.3.3.2 Prétraitement des données . . . . .  | 83 |
| V.3.3.3 Classifieurs et réglage des hyperparamètres . . . . .                                  | 85 |
| V.3.3.4 Évaluation des performances . . . . .  | 89 |
| V.3.3.5 Méta-évaluation . . . . .  | 90 |
| V.3.3.6 Meilleur classifieur . . . . .   | 90 |
| V.4 Résultats . . . . .  | 90 |
| V.4.1 Modèles de prédiction de l'aggravation de l'EDSS . . . . .                               | 90 |
| V.4.2 Modèles de prédiction de l'aggravation des lésions T2 . . . . .                          | 91 |
| Discussion . . . . .   | 92 |

## Introduction

Notre expérience de l'acceptation du CDSS MS Vista (décrit dans le chapitre précédent) par les médecins, notamment dans la mesure où il est basé sur des facteurs utilisés et catégorisés par les médecins experts, nous a conduit à une nouvelle approche hybride pour promouvoir l'adoption des systèmes basés sur le ML dans la pratique médicale.

En effet, la flexibilité et la capacité des algorithmes d'apprentissage automatique (ML-*Machine Learning*), ainsi que la puissance de calcul et de stockage disponible, sont des facteurs privilégiés pour faire progresser la médecine de précision. Les algorithmes de ML sont capables de détecter des différences et des corrélations au sein des données pour faire des prédictions dans des contextes de soins complexes. Cependant, les risques potentiels liés aux données / algorithmes, leur effet sur la sécurité des patients et des médecins, etc. ont créé des obstacles à l'intégration du ML dans la pratique clinique.

Dans ce chapitre, je présente notre approche d'hybridation des intelligences humaine et artificielle, ainsi que les résultats de sa comparaison avec la méthode conventionnelle de création de modèles de prédiction. Ce travail a été publié dans le journal *Sensors* en tant qu'article de recherche du numéro spécial *Human-Machine Intelligence Hybridization : Challenges, Approaches, Applications* par Ed-Driouch et al. [1].

### V.1 État de l'art

L'hybridation humain-machine s'inscrit dans une perspective symbiotique de l'humain et du système dans le courant de pensée de l'intelligence augmentée de C. Engelbart [156]. La fusion de l'intelligence humaine et de l'intelligence artificielle est une forme de collaboration humain-machine pour s'améliorer mutuellement et faire émerger le plus grand potentiel des deux considérées ensemble [96]. Ainsi, l'IA peut fournir un soutien à la prise de décision humaine, ou les humains peuvent fournir des indications au processus de l'apprentissage pour soutenir les fonctions de l'IA.

Les applications du ML pour prédire la progression de la SEP ont été discutées dans des publications récentes par S. Denissen et al. [157] et L. Hone et al. [158]. Seuls A. Tacchella et al. [159] ont mené une étude sur la prédiction de la progression de la SEP basée sur la collaboration de l'intelligence humaine et des algorithmes de ML. Sur la base des

dossiers cliniques recueillis au cours de 527 visites de 84 patients et de l'algorithme Random Forest, ils ont créé un modèle permettant de prédire la transition des patients de la SEP-RR à la SEP-PS après 180, 360 et 720 jours de leurs visites. L'hybridation des intelligences humaine et artificielle dans cette étude intervient dans la phase de prédiction. Les probabilités de prédiction du modèle de ML ont été combinées avec des probabilités faites par des étudiants en médecine sur la base des dossiers cliniques des patients étudiés. Une pondération en fonction de la cohérence des prédictions humaines ou algorithmiques a été effectuée pour obtenir la probabilité de la conversion du patient de la SEP-RR à la SEP-PS dans les 180, 360 et 720 jours. Sur ces 3 périodes, des augmentations de 1.5%, 2.4% et 1.7%, respectivement, du score AUC ont été obtenues par la prédiction hybride comparée au modèle de prédiction seul. Ces résultats sont néanmoins limités par l'absence de prise en compte de l'expérience et du raisonnement des médecins experts de la SEP. Il est essentiel de considérer les connaissances de l'expert humain du domaine en tant que composante principale du processus d'apprentissage de l'algorithme [160] avant de se demander si le modèle créé est valable pour la décision. La qualité et l'utilité clinique des données utilisées ne sont pas claires. Cependant, ces éléments sont importants pour la compréhension et peuvent conduire à des découvertes très importantes. Le modèle n'est renseigné que par l'algorithme qui a été utilisé pour le créer, le processus d'apprentissage demeure une boîte noire. Ces limites ont été prises en compte dans le développement de notre méthode de prédiction hybride.

## V.2 Hypothèse

Nous proposons de combiner le raisonnement intuitif des professionnelles de santé avec la puissance de calcul disponible. Cela pourrait améliorer la qualité des résultats d'une part, en fusionnant les intelligences humaine et artificielle, et gagner la confiance des médecins d'autre part, en les rassurant par l'intégration de leur expertise dans le processus de prédiction. Dans le contexte de la SEP, pour prédire l'évolution de la maladie, la combinaison de l'intelligence des médecins experts et le modèle de prédiction devrait permettre un pronostic plus précis que le clinicien seul ou des algorithmes de prédiction seuls. En effet, l'utilisation de grandes masses de données des études cliniques sur l'efficacité des médicaments a montré son intérêt [141]. Cependant, l'utilisation d'outils

de décision automatiques dans les systèmes critiques a démontré des insuffisances [161], l'hétérogénéité des facteurs prédictifs, qui diffèrent d'un patient à l'autre, complique le pronostic [142], [162], [163], et la variabilité dans la composition de la SEP rend les données d'entrée et de sortie difficiles à sélectionner et à prédire au milieu des différentes méthodes d'IA.

### **V.3 Nouvelle approche de l'intelligence hybride**

Nous avons proposé une nouvelle approche de prédiction de l'évolution de la SEP basée sur le développement de modèles prédictifs hybrides exploitant la puissance de calcul, des données cliniques de bonne qualité et l'interaction avec des médecins experts [3]. Cette approche vise à fusionner les intelligences humaines et artificielles afin de concevoir des modèles hybrides de ML pertinents et compréhensibles pour les humains. Pour valider cette approche de collaboration humain-algorithme, nous avons comparé la méthode conventionnelle de création de modèles de prédiction avec notre méthode hybride. Dans cette section, je présente le contexte de l'approche proposée, le pipeline de création de modèles de prédiction qui a été développé pour créer des modèles basés sur la méthode hybride et la méthode conventionnelle, ainsi que les résultats de la comparaison de celles-ci.

#### **V.3.1 Collaboration médecin-algorithme**

Notre approche de prédiction hybride vise à intégrer l'expérience et les connaissances des médecins experts dans le processus de prédiction, en particulier dans la phase de préparation des données pour concevoir des modèles de prédiction basée sur la collaboration médecin-algorithme (voir Figure V.1). Nous nous sommes particulièrement intéressés à la prédiction de l'évolution de la maladie faite par les médecins grâce à des mesures (par exemple, l'IRM), des données médicales (par exemple, le calcul de l'EDSS), et des observations directes du patient (consultations). L'analyse d'activité réalisée pour le développement du CDSS MS Vista décrit dans le chapitre précédent a servi de base à notre réflexion. Notre approche hybride consiste à utiliser les catégories définies avec les neurologues pour la contextualisation des patients atteints de SEP (voir section IV.4.2.1.

Contextualisation : POI / POR du Chapitre IV) pour transformer les données d'entraînement conformément au raisonnement du médecin expert. Au lieu d'utiliser les valeurs d'un facteur du patient et de le catégoriser automatiquement, nous proposons d'utiliser plutôt le groupe équivalent basé sur le raisonnement des médecins. Prenons un exemple de la Figure V.1, pour le facteur « âge au début de la maladie », le patient avait 30 ans, selon le raisonnement des médecins ce patient appartient au groupe A2 (âge compris entre [20-40] ans). Par conséquent, plutôt que d'utiliser la valeur brute de ce facteur, nous utiliserions le groupe équivalent. Le même principe de transformation des données est appliqué pour les autres facteurs de tous les patients de la base de données. Les données transformées seront ensuite utilisées pour créer un modèle prédictif des indicateurs de l'activité de la maladie. Les catégories de facteurs utilisées pour la transformation des données sont décrites dans le Tableau IV.2 du Chapitre IV.

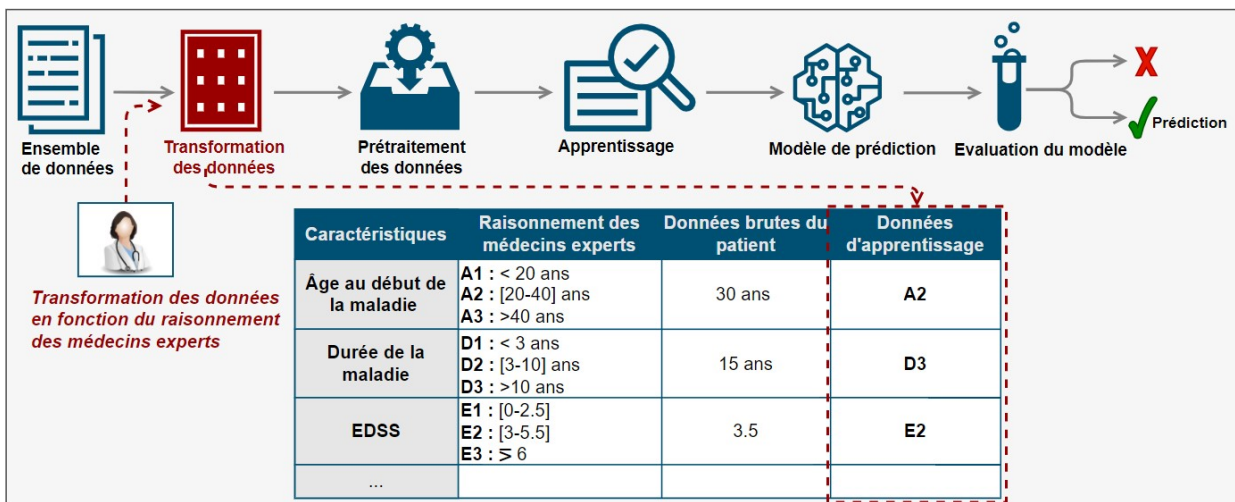


FIGURE V.1 – L'intégration du raisonnement des médecins dans le processus de création du modèle de prédiction (source Ed-Driouch et al.[1])

*La transformation des données (bloc rouge) avant de procéder à l'apprentissage (voir les détails dans la Figure V.2) : au lieu d'utiliser les valeurs d'un facteur du patient, nous utilisons le groupe correspondant au raisonnement des médecins.*

### V.3.2 Évaluation de l'approche d'intelligence hybride médecin-algorithme

Pour évaluer l'approche collaborative proposée, nous avons utilisé les données de patients atteints de SEP issues de l'essai clinique (NCT00906399) (voir la section IV.3.



Enjeux des données du Chapitre IV pour plus d'information sur les données) afin créer des modèles prédictifs des indicateurs suivants de l'activité de la SEP :

- EDSS : progression du handicap
- Nouvelles lésions : apparition de nouvelles lésions sur l'IRM

Pour chacun de ces deux facteurs, nous avons appliqué le pipeline de création et de sélection de modèle de prédiction (Figure V.2) décrit ci-dessous, pour concevoir trois modèles de prédiction basés sur différentes méthodes de prétraitements de données :

- *All features* : modèle créé en utilisant tous les facteurs existants dans la base de données, soit la méthode conventionnelle.
- *Physician features* : modèle créé uniquement à partir des facteurs utilisés par les médecins de la SEP pour prendre une décision médicale, soit une première étape dans la prise en compte des compétences des experts humains.
- *Physician features & categories* : modèle créé sur la base des catégories de facteurs utilisés par les neurologues pour prendre une décision médicale. Il s'agit d'une prise en compte plus approfondie du raisonnement des experts médicaux, celui qu'ils mettent en œuvre pour catégoriser les patients afin d'établir le pronostic. Au lieu d'utiliser les données brutes des patients pour entraîner le modèle (un patient âgé de 32 ans au début de la maladie), la catégorie équivalente au raisonnement des médecins est utilisée (un patient dans la catégorie [20-40]).

### V.3.3 Classes déséquilibrées : pipeline pour la création de modèles de prédiction

Le déséquilibre des classes est un problème typique de la recherche dans la SEP, car la maladie évolue lentement. Par conséquent, les ensembles de données contiennent souvent beaucoup plus d'enregistrements liés à la classe « pas d'aggravation » qu'à la classe « aggravation » (voir Tableau V.1). Ceci peut conduire à des modèles biaisés si le problème n'est pas traité correctement. Pour réduire cet effet de biais, nous avons utilisé les techniques suivantes :

- L'utilisation de différents algorithmes de prédiction

- La pondération
- L'optimisation des paramètres des algorithmes

Ces techniques ont été employées dans le pipeline de création et de sélection des modèles de prédiction (voir la Figure V.2) décrit ci-dessous.

TABLEAU V.1 – Proportion de chaque classe pour les facteurs à prédire

| Indicateur | Classes           |             |
|------------|-------------------|-------------|
|            | Pas d'aggravation | Aggravation |
| EDSS       | 728 (~93%)        | 51 (~7%)    |
| Lésions T2 | 394 (~51%)        | 385 (~49%)  |

### V.3.3.1 Séparation des données

La première étape du pipeline consiste à séparer les facteurs (*features*) qui seront utilisés par le modèle pour prédire la cible (*target*) du modèle créé (voir l'étape 1 de la figure V.2), c'est-à-dire l'indicateur de l'évolution de la maladie. Nous avons utilisé les données des patients au début de l'essai clinique pour créer des modèles prédisant l'évolution à un an des deux indicateurs de l'activité de la SEP décrits dans le Tableau V.2.

TABLEAU V.2 – Explication de l'aggravation des indicateurs à prédire

| Indicateur | Aggravation  |
|------------|--|
| EDSS       | Progression d'un point pour un EDSS compris entre [1-5.5]<br>Progression de 0.5 point pour un EDSS $\geq 6$<br>Progression de 1.5 point pour un EDSS = 0 |
| Lésions T2 | Apparition de nouvelles lésions sur l'IRM  |

### V.3.3.2 Prétraitement des données

Les estimateurs de ML peuvent fournir des résultats non conformes si les facteurs individuels ne correspondent pas plus ou moins à des données normalement distribuées. Nous avons remarqué que les facteurs de notre ensemble de données couvrent différents

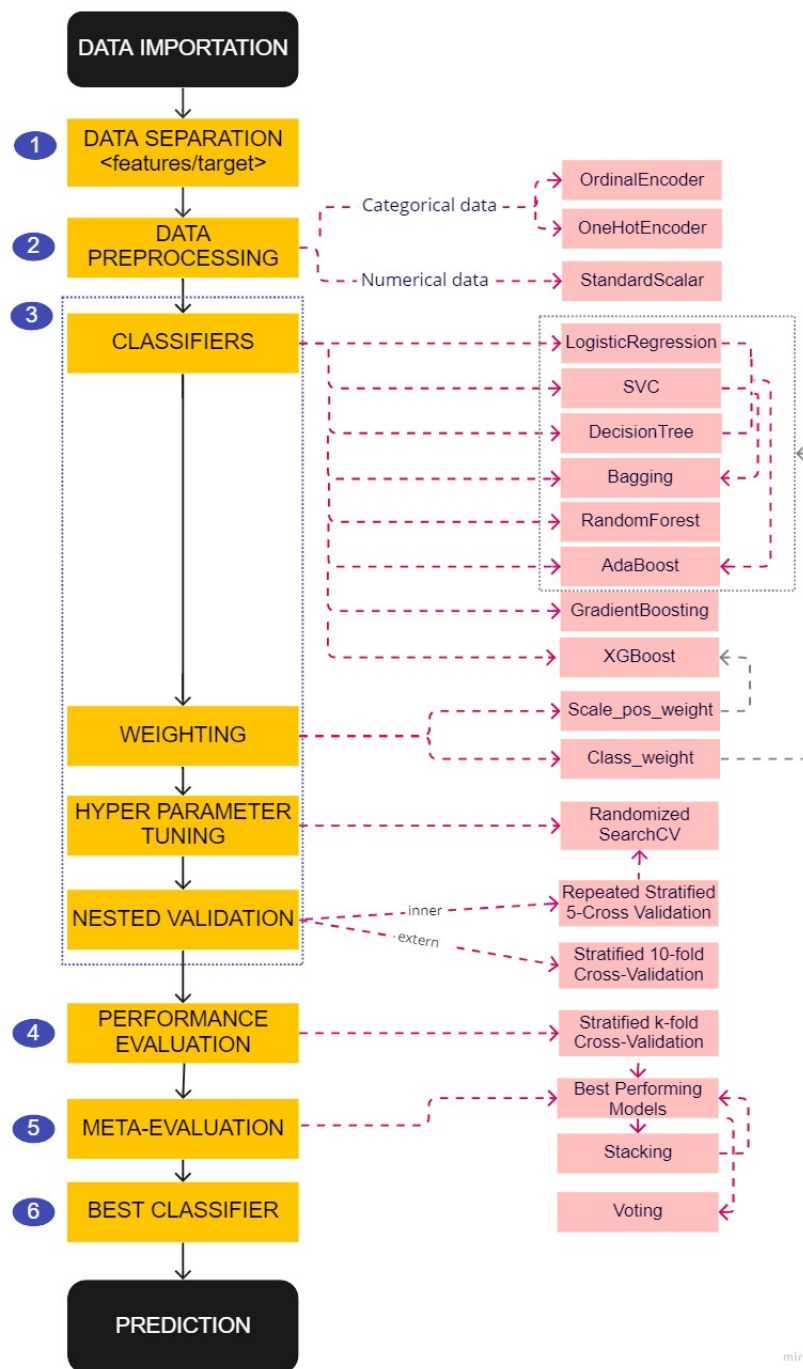


FIGURE V.2 – Pipeline pour la création et la sélection des modèles de prédiction (source Ed-Driouch et al.[1])

Les numéros dans les cercles en bleu renvoient aux différentes sous-sections de la section « V.2.3. Classes déséquilibrées : pipeline pour la création de modèles de prédiction », qui représente les étapes du pipeline indiquées dans les cases jaunes. Les flèches noires indiquent la séquence des étapes. Les flèches roses en pointillés indiquent les différentes techniques utilisées à chaque étape. Les flèches grises en pointillés indiquent les algorithmes auxquels chaque paramètre de pondération a été appliqué.

types de distributions (voir en Annexe A, le Tableau A.1 et les Figures A.1-A.4). Ainsi, nous avons utilisé un transformateur pour normaliser nos données (voir l'étape 2 de la Figure V.2). Nous avons employé *StandardScalar*, qui décale et met à l'échelle chaque facteur individuellement afin qu'ils aient tous une moyenne de 0 et un écart-type unitaire.

Concernant les données catégorielles, deux transformateurs différents ont été utilisés. Le choix du transformateur dépend du modèle sous-jacent et du type de catégories (nominales/ordinales). En général, *OneHotEncoder* est la stratégie de codage utilisée lorsque les modèles sont linéaires. Elle empêche les modèles de faire une fausse hypothèse sur l'ordre des catégories en créant autant de nouvelles colonnes qu'il y a de catégories possibles. Elle consiste à prendre une variable catégorielle et à la transformer en plusieurs variables binaires. Chaque variable binaire correspond à une valeur unique de la variable catégorielle, et prend la valeur 1 si la valeur d'origine correspondante est présente, et 0 sinon. Par ailleurs, la technique de transformation *OrdinalEncoder* est utilisée avec les modèles basés sur des arbres. Elle assigne une valeur numérique unique à chaque valeur catégorielle. Les valeurs numériques sont assignées en fonction de l'ordre des catégories, ce qui signifie que les catégories qui apparaissent en premier reçoivent une valeur numérique plus faible, et ainsi de suite (par exemple,  $0 < 1 < 2$ ).

### V.3.3.3 Classifieurs et réglage des hyperparamètres

Pour réduire le biais du déséquilibre des classes, nous avons utilisé les méthodes suivantes (voir l'étape 3 de la Figure V.2) :

1. Différents modèles de prédiction : les performances des modèles de prédiction diffèrent en fonction de la distribution des données et des classes à prédire. Nous avons entraîné différents algorithmes afin de choisir le modèle de prédiction le plus performant parmi les modèles créés. Trois algorithmes classiques ont été utilisés comme baseline [164] : *Support Vector Classifier* (SVC), *Logistic regression*, *Decision tree*, et sept algorithmes de la technique d'apprentissage ensembliste (*Ensemble Learning*) [165] (Plus de détails sur les algorithmes utilisés sont fournis dans Annexe A, Tableau A.2). Cette technique consiste à entraîner plusieurs algorithmes pour ensuite considérer l'ensemble de leurs prédictions. Différentes méthodes de l'apprentissage ensembliste sont disponibles pour améliorer la stabilité et la précision des algorithmes

de ML, tout en permettant de réduire la variance de l'estimateur et d'éviter le sur-apprentissage. Nous avons utilisé les méthodes suivantes :

- *Bagging (Bootstrap AGGregatING)* : cette méthode consiste à créer plusieurs entités (estimateur) d'un même modèle (par exemple, le modèle de baseline) et à entraîner chacune de ces entités sur une portion aléatoire de l'ensemble de donnée. Ceci est réalisé en utilisant le *Bootstrapping* qui est une technique d'échantillonnage qui consiste à remplacer après chaque tirage au sort les données qui ont été sélectionnées dans notre ensemble de données. De cette façon, nous obtenons des modèles diversifiés puisqu'ils n'ont pas tous été formés sur les mêmes données. Cependant, ils partagent des connaissances communes. Ensuite, les résultats sont regroupés pour choisir par vote la bonne prédiction. Nous avons utilisé l'algorithme *Bagging* trois fois avec les modèles de baseline comme estimateurs. Nous avons également utilisé l'algorithme *RandomForest* qui utilise la technique du *Bagging* avec toujours comme estimateur *Decision-Tree* et une sélection aléatoire des *features*.
- *Boosting* : cette méthode consiste à entraîner plusieurs modèles (en série) relativement faibles, en demandant à chaque modèle d'essayer de corriger les erreurs commises par son prédécesseur. Au final, on obtient un ensemble de modèles complémentaires dans lesquels les faiblesses des uns sont compensées par les forces des autres. Nous avons utilisé les algorithmes *Adaboost* (3 fois avec les modèles de baseline comme estimateurs), *GradientBoosting* et *XG-Boost (Extreme Gradient Boosting)*.
- *Stacking* : il s'agit d'une méthode d'ensemble où les modèles sont combinés à l'aide d'un autre algorithme. L'idée de base est d'entraîner les algorithmes avec un ensemble de données d'entraînement, puis d'utiliser leur sortie comme entrée de l'algorithme combineur.
- *Voting* : cette méthode consiste à créer plusieurs modèles de classification en utilisant un ensemble de données d'entraînement. Nous avons utilisé le *soft voting*, où chaque classifieur individuel fournit une valeur de probabilité qu'un point de données spécifique appartient à une classe particulière. Les prédictions peuvent être pondérées par l'importance de chaque classifieur en fonction

de sa performance (à travers le réglage du paramètre *weight* de l'algorithme) et sont ensuite additionnées. La classe avec la plus grande somme de probabilités (pondérées) remporte le vote.

2. Pondération : pour les algorithmes de baseline, *Bagging*, *RandomForest*, et *Adaboost*, nous avons utilisé le paramètre *class\_weight* avec la valeur *balanced* pour pénaliser les erreurs sur la classe minoritaire par un poids proportionnel à sa sous-représentation. Pour l'algorithme *XGBoost*, nous avons utilisé le paramètre *scale\_pos\_weight* pour contrôler l'équilibre des poids positifs et négatifs.
3. Optimisation des paramètres (*Hyperparameter tuning*) : l'optimisation des paramètres du modèle peut produire de meilleurs résultats que l'utilisation des paramètres par défaut pour classifier des ensembles de données non équilibrés [166]. Nous avons appliqué 30 itérations de l'approche *RandomizedSearchCV* [167] pour trouver les meilleurs paramètres pour chaque algorithme (à l'exception des algorithmes *Stacking* et *Voting*, car nous les avons utilisés pour effectuer une méta-validation) en fonction de notre ensemble de données.
4. Validation croisée imbriquée (*Nested cross-validation*) : nous avons appliqué la validation croisée imbriquée qui offre un moyen de réduire le biais dans le réglage combiné des paramètres et la sélection du modèle. Plus précisément, nous avons effectué une validation croisée externe à 10 plis qui divise le jeu de données en dix plis stratifiés. Cela permet de créer des plis avec approximativement les mêmes proportions de cas de chaque classe que l'ensemble de données complet, ce qui permet d'éviter qu'un jeu d'entraînement ne contienne que des cas positifs et que l'ensemble de test correspondant ne contienne que des cas négatifs. Chacun des dix plis est ensuite considéré comme une donnée de test, tandis que les autres plis constituent les données d'entraînement. Pour chaque ensemble d'apprentissage, nous avons appliqué une validation croisée stratifiée à 5 plis, répétée 4 fois, pour sélectionner les paramètres avec la méthode *RandomizedSearchCV*. Ceci a été réalisé sur la base de la valeur la plus élevée du score *balanced accuracy* (voir Tableau V.3), c'est-à-dire la moyenne de la proportion de réponses correctes de chaque classe considérée individuellement. Par conséquent, pour chaque modèle, nous avons obtenu 10 ensembles de paramètres différents. Nous nous sommes limités aux paramètres du modèle as-

socié à la valeur maximale du score *balanced accuracy*.

TABLEAU V.3 – Description des scores utilisés pour mesurer les performances des modèles de prédiction.

| Score  | Définition  | Formule   |
|--|---|---|
| VP (Vrai Positif)                                | Nombre de cas où le modèle a correctement prédit une observation positive.  | -   |
| VN (Vrai Négatif)                                | Nombre de cas où le modèle a correctement prédit une observation négative.  | -   |
| FP (Faux Positif)                                | Nombre de cas où le modèle a incorrectement prédit une observation positive.  | -   |
| FN (Faux Négatif)                                | Nombre de cas où le modèle a incorrectement prédit une observation négative.  | -   |
| Précision  | Proportion de prédictions positives correctes.  | $\text{Précision} = \text{VP} / (\text{VP} + \text{FP})$  |
| Rappel   | Proportion d'observations positives correctement identifiées.   | $\text{Rappel} = \text{VP} / (\text{VP} + \text{FN})$   |
| ROC ( <i>Receiver Operating Characteristic</i> ) | Courbe représentant la relation entre le taux de faux positifs et le taux de vrais positifs pour différents seuils de classification.   | -   |
| AUC ( <i>Area Under the ROC</i> )                | Représente la proportion de la surface totale sous la courbe ROC d'un modèle. Plus précisément, l'AUC mesure la capacité d'un modèle à distinguer entre les classes positives et négatives. | Se calcule en intégrant la courbe ROC. Plus la surface totale située en dessous de la courbe est grande, plus le modèle est performant. |

TABLEAU A.2 - Description des scores utilisés pour mesurer les performances des modèles de prédiction. (Suite)

| Score             | Définition   | Formule  |
|-------------------|--|--|
| F1 Score          | Indicateur de la performance d'un modèle de classification binaire qui mesure le compromis entre la précision et le rappel.  | $F1 = 2 * (\text{Précision} * \text{Rappel}) / (\text{Précision} + \text{Rappel})$   |
| Balanced Accuracy | Moyenne pondérée des taux de vrais positifs et de vrais négatifs, où les poids sont égaux aux proportions d'échantillons positifs et négatifs dans les données d'entraînement. | $\text{Balanced Accuracy} = (\text{TPR} + \text{TNR}) / 2$<br>* TPR (taux de vrais positifs) est le nombre de positifs correctement classés divisé par le nombre total de vrais positifs.<br>* TNR (taux de vrais négatifs) est le nombre de négatifs correctement classés divisé par le nombre total de vrais négatifs. |

*Seuls les scores AUC, F1 et Balanced Accuracy ont été utilisés dans cette étude. Les autres scores sont fournis pour clarifier le calcul des scores utilisés.*

#### V.3.3.4 Évaluation des performances

Les modèles sont réévalués (voir l'étape 4 de la Figure V.2) avec une validation croisée stratifiée à 5 plis. Ils sont ensuite filtrés en fonction de la moyenne et l'écart-type du score AUC (l'aire sous la courbe ROC (*Receiver Operating Characteristic*)) [168] (voir Tableau V.3).



### V.3.3.5 Méta-évaluation

Une méta-évaluation des modèles a été appliquée (voir l'étape 5 de la Figure V.2). Les modèles obtenus par le filtrage précédent ont été utilisés comme estimateurs pour l'algorithme *Stacking*. Un méta-filtrage similaire au précédent a été appliqué pour décider de conserver le *Stacking* parmi les modèles les plus performants. L'algorithme *Voting* a également été utilisé avec les modèles obtenus par le méta-filtrage comme estimateurs.

### V.3.3.6 Meilleur classifieur

Le meilleur modèle de prédiction est celui qui présente la valeur moyenne maximale du score AUC et la valeur minimale de son écart-type.

## V.4 Résultats

Pour évaluer la méthode proposée, nous avons utilisé les données de patients sous placebo (pas de traitement) et sous peginterféron bêta-1a avec des doses approuvées par la FDA (*Food and Drug Administration*). Les données de 779 patients au début de l'essai clinique ont été utilisées pour créer un modèle prédictif des indicateurs de l'activité de la SEP un an plus tard. Nous avons appliqué le pipeline de modèles de prédiction décrit ci-dessus pour créer et sélectionner les trois types de modèles : *All features*, *Physician features*, *Physician features & categories* (voir la section V.2.b Évaluation de l'approche d'intelligence hybride médecin-algorithme) pour chaque indicateur. Par la suite, nous avons comparé les performances des trois types de modèles de prédiction de chaque indicateur (voir plus de détails dans Annexe A, les Tableaux A.3-A.8). Enfin, nous avons utilisé les scores AUC et F1 (voir le Tableau V.3) pour évaluer les performances des modèles. Les résultats obtenus sont décrits dans les sous-sections suivantes.

### V.4.1 Modèles de prédiction de l'aggravation de l'EDSS

La Figure V.3 illustre les performances des modèles sélectionnés pour chaque type de modèle créé pour la prédiction de l'aggravation de l'EDSS. Pour le modèle *All features*, nous avons trouvé un score AUC moyen de 0.53 et un score F1 moyen de 0.66. Le modèle

sélectionné pour la méthode *Physician features* a montré une augmentation des performances de +0.03 pour le score AUC et de +0.05 pour le score F1. Une différence plus importante a été constatée en comparant les performances du modèle *All features* à celles du modèle *Physician features & categories*, le score AUC a augmenté de +0.07 et le score F1 de +0.08 avec des valeurs d'écart-type plus faibles (valeurs std dans la Figure V.3).

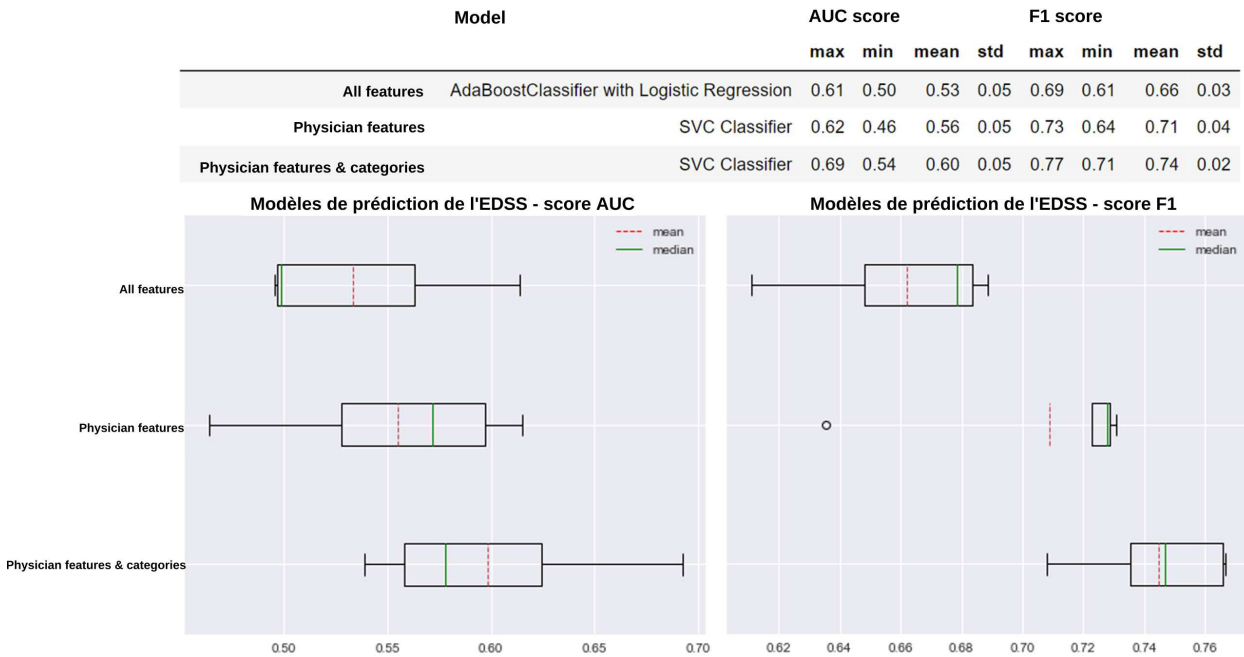


FIGURE V.3 – Modèles prédictifs sélectionnés pour chaque type de modèle de prédiction de l'EDSS avec leurs scores AUC et F1 (source Ed-Driouch et al.[1]).

*Les diagrammes en boîte montrent une description visuelle des performances des modèles de prédiction de l'EDSS sélectionnés pour chaque méthode de prédiction.*

#### V.4.2 Modèles de prédiction de l'aggravation des lésions T2

La Figure V.4 montre les performances des modèles sélectionnés pour chaque type de modèle créé pour la prédiction de l'aggravation des lésions T2. Les scores AUC et F1 du modèle basé sur la méthode *All features* étaient de  $0.72 \pm 0.04$ . Des résultats similaires ont été obtenus pour les modèles basés sur les méthodes *Physician features* et *Physician features & categories* avec des différences de seulement 0.01. L'écart type des scores AUC et F1 du modèle *All features* dépasse de 0.01 l'écart type des scores AUC et F1 obtenus pour le modèle *Physician features*. De plus, les moyennes des scores AUC et F1 du modèle

All features étaient inférieures de 0.01 à celles du modèle *Physician features* et du modèle *Physician features & categories*.

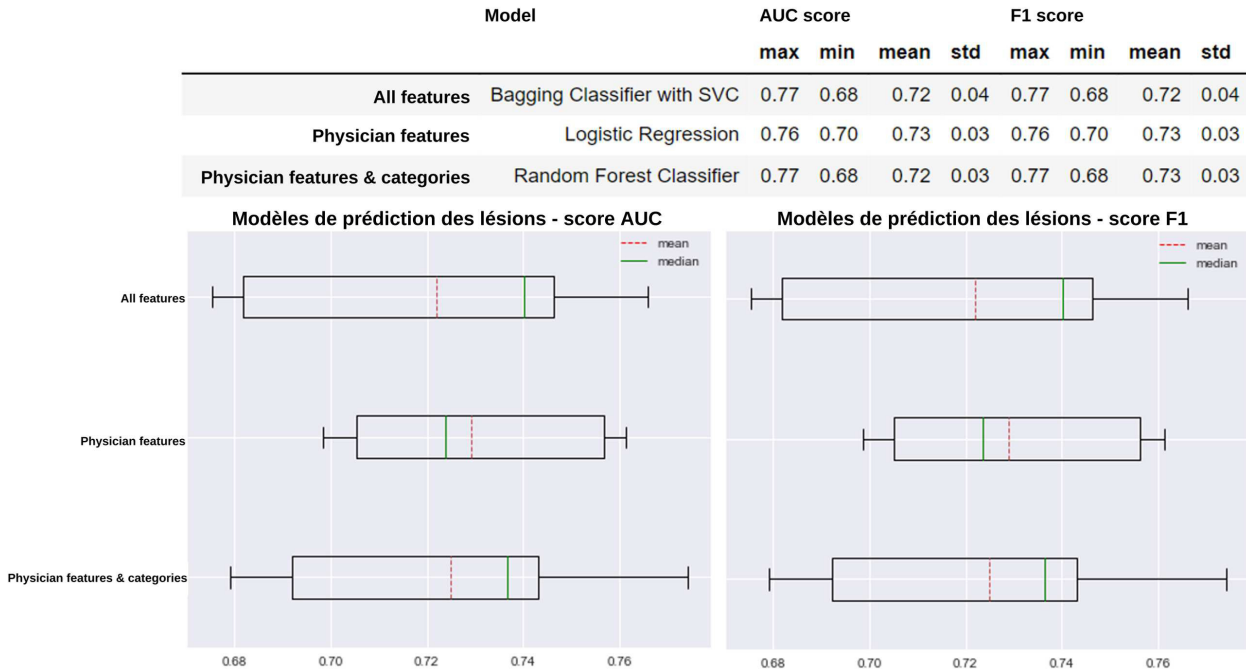


FIGURE V.4 – Modèles prédictifs sélectionnés pour chaque type de modèle de prédiction des lésions T2 avec leurs scores AUC et F1 (source Ed-Driouch et al.[1]).

*Les diagrammes en boîte montrent une description visuelle des performances des modèles de prédiction des nouvelles lésions T2 sélectionnés pour chaque méthode de prédiction.*

## Discussion

Ce chapitre a présenté une nouvelle méthode de développement de modèles prédictifs hybrides qui exploite la puissance des algorithmes de ML, des données de bonne qualité et des interactions avec des médecins experts pour aider à relever les défis de l'adoption du ML dans la pratique médicale. Cette nouvelle méthode d'intelligence hybride est basée sur la collaboration médecin-algorithme. Elle intègre le raisonnement des médecins experts de la SEP dans le processus d'apprentissage des algorithmes de ML. Notre cas d'application concerne la prédiction de la progression de la SEP.

Nous avons conçu notre approche de sorte qu'elle réponde aux limites de l'approche de prédiction hybride déjà développée pour la sclérose en plaques présentée dans la

section « État de l'art » de ce chapitre. Pour remédier aux problèmes liés aux données, nous avons utilisé des données cliniques de bonne qualité adaptées à l'objectif de création du modèle de prédiction. En effet, pour prédire la progression de la SEP en fonction de l'efficacité du traitement, nous avons utilisé les données d'un essai clinique de phase III évaluant l'efficacité d'un traitement de SEP par rapport au placebo. Pour les patients sous traitement, nous n'avons utilisé que les données des patients recevant des doses approuvées par la FDA. Ainsi, nous sommes sûrs d'entraîner un modèle avec des données qui reflètent son utilisation clinique. Pour faciliter la compréhension du processus d'apprentissage et guider la décision du modèle en fonction de la pratique médicale, nous nous sommes appuyés sur l'expérience et le raisonnement des médecins experts. Au lieu d'utiliser les données brutes des patients pour entraîner le modèle, nous avons utilisé les catégories correspondant au raisonnement des médecins concernant chaque facteur. Les données catégorisées sont ensuite utilisées pour entraîner l'algorithme de ML afin de créer un modèle de prédiction fondé sur le processus de décision de la pratique médicale.

Pour évaluer l'approche proposée, nous avons développé un pipeline de modèles de prédiction détaillés qui traite le problème du déséquilibre des classes, qui est un problème typique de la SEP. Nous avons appliqué ce pipeline pour créer des modèles de prédiction de l'EDSS et de l'aggravation des lésions T2. Pour chaque indicateur, nous avons sélectionné trois types de modèles de prédiction : (1) modèle basé sur toutes les facteurs présents dans l'ensemble de données, (2) modèle basé uniquement sur les facteurs utilisés par le médecin expert pour prendre une décision, et (3) modèle basé sur les facteurs et leurs catégories basées sur le raisonnement du médecin. Nous avons comparé les performances des modèles obtenus pour chacun des deux indicateurs, EDSS et lésions. Malgré le grand déséquilibre entre les classes de l'EDSS (pas d'aggravation : 728 patients vs. aggravation : 51 patients), une différence significative a été trouvée entre les différents types de modèles, en particulier entre le modèle (3) et le modèle (1). Le score AUC du modèle (3) est supérieur de 4% à celui du modèle (2) et de 7% à celui du modèle (1). Le score F1 du modèle (3) est supérieur de 3% à celui du modèle (2) et de 8% à celui du modèle (1). La différence entre les trois types de modèles de prédiction montre que la performance du modèle de prédiction (3) est la meilleure, ce qui montre l'utilité de notre méthode de prédiction hybride. Pour la prédiction de l'aggravation des lésions, nous avons constaté que les trois types de modèles ont approximativement les mêmes performances (avec des

différences de 1%). Cela confirme qu'il est suffisant de s'appuyer sur le raisonnement du médecin expert pour créer des modèles de prédiction. En outre, les modèles qui incluent le raisonnement du médecin sont plus rassurants, car ils orientent la décision du modèle de prédiction en fonction de la pratique médicale (l'expérience du médecin expert).

Bien que le modèle basé sur notre approche hybride de la prédiction de l'EDSS soit meilleur que l'approche conventionnelle, il n'est pas encore en mesure de servir à la prise de décision médicale. La raison peut en être l'absence de certains facteurs pathologiques dans l'ensemble de données ou le manque d'utilisation de données à long terme. D'autres facteurs tels que les biomarqueurs prédictifs de la SEP (le neurofilament sérique-chaîne légère, la protéine acide fibrillaire gliale, etc.), les effets (secondaires) à long terme des options thérapeutiques, les déterminants sociaux de la santé, l'état émotionnel des patients, etc. peuvent être utilisés pour entraîner les algorithmes de ML. L'utilisation complémentaire de ces facteurs a le potentiel d'améliorer la précision des modèles de prédiction. Des facteurs tels que les déterminants sociaux de la santé peuvent nécessiter l'intégration d'experts humains d'autres domaines (sociologie, psychologie, etc.) dans le processus de sélection et de catégorisation des facteurs. Ainsi, notre approche pourrait s'étendre à des collaborations étroites avec de tels experts. Outre l'amélioration de la pertinence du modèle de prédiction, cela pourrait également ouvrir de nouvelles voies de recherche dans les autres disciplines.

Le concept de cette approche hybride a été repris dans un autre projet visant à anonymiser les données des patients. Au lieu d'appliquer l'algorithme d'anonymisation sur les données brutes pour créer des avatars (données synthétiques anonymes issues d'une simulation mathématique basée sur les données réelles des patients), les données transformées basées sur notre approche hybride ont été utilisées. L'anonymisation hybride a montré une augmentation des résultats de l'anonymisation et une incapacité à ré-identifier les patients par leurs avatars, étant donné qu'ils ne sont pas créés de façon directe à partir des données réelles.

L'hybridation des intelligences humaine et artificielle est une voie prometteuse pour surmonter les obstacles qui limitent la mise en pratique des systèmes basés sur l'IA dans les systèmes de santé et promouvoir un avenir sûr et révolutionnaire pour la médecine. La méthode proposée est destinée à permettre des applications fiables et acceptables des algorithmes de ML pour soutenir la personnalisation de la prise de décision clinique.

# CONCLUSION GÉNÉRALE ET PERSPECTIVES

Cette thèse souligne l'importance des interactions humain-machine dans le développement de méthodes d'aide à la décision. L'objectif principal est d'établir les bases d'une collaboration humain-machine fonctionnelle, afin de développer des solutions numériques d'aide à la décision et de promouvoir leur adoption dans les services de santé. Notre cas d'application concerne l'utilisation des données de patients atteints de maladies chroniques pour soutenir les pratiques de prise de décision médicale individualisée. L'utilisation de telles données, dont l'importance est vitale pour l'aide à la décision, nécessite une attention particulière à leur traitement et à leur représentation. Des données mal traitées ou mal représentées peuvent être mal comprises et donc mal interprétées, ce qui peut influencer la prise de décision et créer des préoccupations majeures [169]. Il est donc essentiel d'intégrer les professionnels de santé dans le processus de traitement des données et de leur accorder un contrôle direct sur leur utilisation, pour les sélectionner, les évaluer et déterminer si elles sont appropriées pour l'aide à la décision.

Dans cette perspective, nous avons développé un CDSS (*Clinical Decision Support System*) de médecine de précision, appelé MS Vista, pour aider dans l'individualisation de la décision thérapeutique dans la SEP (Sclérose En Plaques). Le CDSS est basé sur l'utilisation de données cliniques de bonne qualité et leur filtrage, sur la base de facteurs et de catégories validées dans la littérature et par des médecins experts, pour contextualiser les patients atteints de SEP. Il permet un calcul à la demande pour obtenir les informations nécessaires pour aider à prendre les bonnes décisions au bon moment, tout en étant transparent en termes de données utilisées. Il permet au clinicien d'examiner la qualité de la base de données dans son ensemble et la qualité des données utilisées pour contextualiser chaque patient et de décider si elles sont conformes et appropriées pour la prise de décision. Grâce aux projections de données, MS Vista peut aider le médecin à établir des prévisions sur l'évolution potentielle des principaux indicateurs de l'activité de la SEP, et à choisir le traitement le plus approprié pour chaque patient. Il illustre l'impact des facteurs de ce dernier sur la progression de sa maladie et les avantages potentiels de différentes options thérapeutiques. Par ailleurs, l'utilité de MS Vista dans la décision thérapeutique a été confirmée par des neurologues experts. Cependant, elle est limitée

par le manque de certaines informations dans les données. MS Vista est également conçu pour aider le médecin dans sa communication avec le patient, grâce à son mode dual de visualisation. Le premier fournit au médecin des détails sur les données et les concepts utilisés. Le second est destiné à enrichir la communication médecin-patient sur les effets potentiels des différentes options thérapeutiques. Le CDSS MS Vista est une preuve de concept qui a aidé à obtenir un financement de l'ANR [ANR-21-RHUS-0014] pour le projet PRIMUS. Ce dernier vise à enrichir le CDSS MS Vista avec plus de données sur d'autres traitements (plus largement utilisés) et formes de SEP. Ceci grâce à l'accès à des cohortes longitudinales complètes de haute qualité et à des collaborations engagées avec des entreprises pharmaceutiques. Cette démarche permettra une présentation plus personnalisée de l'efficacité à long terme des médicaments et de la progression potentielle de la maladie. Dans cette optique, le CDSS peut également contribuer à améliorer la conformité réglementaire en garantissant que les décisions de traitement sont prises conformément aux normes de qualité et de sécurité établies, telles que la norme ISO 13485 (norme internationale définissant les exigences relatives aux systèmes de gestion de la qualité dans le domaine médical) et la norme ISO 27001 (norme internationale définissant les exigences relatives à la gestion de la sécurité de l'information), et ainsi influencer positivement le marquage CE (certification obligatoire pour les dispositifs médicaux commercialisés dans l'Union Européenne).

Les retours positifs obtenus de la part des médecins concernant MS Vista, notamment du fait qu'il est basé sur des facteurs utilisés et catégorisés par eux, nous a conduit à concevoir ensuite une nouvelle approche hybride pour la création de modèles de prédiction. L'objectif était d'aider à relever les défis de l'adoption de systèmes basés sur l'IA dans les soins de santé. Il s'agit de créer des modèles hybrides pour prédire l'évolution de la SEP en intégrant le raisonnement des médecins experts dans le processus d'apprentissage des algorithmes de ML. Au lieu d'utiliser les données brutes du patient pour entraîner l'algorithme (un patient âgé de 30 ans au début de la maladie), la catégorie équivalente basée sur le raisonnement des médecins est utilisée (un patient dans la catégorie [20-40]). Les résultats obtenus montrent que notre approche hybride est capable de produire des modèles plus performants que l'approche conventionnelle. Notre approche hybride a été utilisée pour aider à anonymiser les données des patients. Au lieu d'appliquer l'algorithme d'anonymisation aux données brutes, les données ont été transformées

sur la base du concept de notre approche hybride. Les résultats ont montré une augmentation des performances avec une incapacité à ré-identifier les patients, étant donné que l'anonymisation n'est pas directement dérivée de données réelles.

Par la suite, plusieurs perspectives peuvent être envisagées pour approfondir le travail effectué dans cette thèse. D'un côté, il serait intéressant de considérer l'effet potentiel à long terme du CDSS MS Vista sur la relation médecin-patient [18]. Des évaluations au fil du temps pourraient conduire à de nouvelles fonctionnalités afin de garantir une durée de vie de MS Vista plus importante. Ces évaluations pourraient également contribuer à surveiller les changements réglementaires et les exigences de conformité, en garantissant que le CDSS est continuellement mis à jour pour répondre aux nouvelles exigences en matière de réglementation et de sécurité. Des analyses de son utilisation et des entretiens réguliers avec les utilisateurs permettraient aussi de consolider le CDSS. Il serait intéressant d'appliquer les concepts employés dans MS Vista pour l'aide à la décision dans la SEP pour aider à la gestion d'autres maladies chroniques. D'un autre côté, l'application de l'approche de prédiction hybride proposée dans cette thèse peut jouer un rôle très important dans divers aspects. Nichol et al. [170] ont mené une étude sur la perception des développeurs de ML destinés aux soins de santé vis-à-vis de la responsabilité des préjudices causés par leur travail. L'analyse d'entretiens avec 40 développeurs a soulevé des questions indiquant la présence d'une composante morale dans ces préjudices. Les développeurs ont indiqué que ces préjudices pouvaient être causés par les actions des développeurs et que ceux-ci ou d'autres personnes pouvaient avoir la responsabilité de les atténuer. L'une des conséquences d'usage de notre approche peut consister en l'atténuation de ces préjudices. Elle pourrait également contribuer à améliorer le positionnement moral des développeurs d'IA pour la santé, en les rassurant que leurs algorithmes sont fondés sur le raisonnement d'experts médicaux. Cette approche hybride est destinée à être appliquée sur d'autres types de données/maladies afin d'aller dans le sens d'applications plus fiables et acceptables des algorithmes de ML. Il pourrait être utile de créer des modèles basés sur des facteurs de la maladie utilisés et catégorisés par des médecins experts (le concept de notre approche hybride), combiné avec des facteurs identifiés comme étant d'intérêt pour la prise de décision médicale mais qui n'ont pas été mis en évidence par des études cliniques (dans le contexte de la SEP, par exemple, l'implication du sel alimentaire dans la pathogenèse de l'auto-immunité du SNC, le stress psycholo-



gique, etc.). La comparaison des performances des modèles créés par cette méthode et la méthode hybride proposée pourrait aboutir à des découvertes sur l'utilité de ces facteurs et à une amélioration de l'approche hybride d'origine. C'est une forme de l'intelligence collective qui peut avoir un effet déterminant dans la prise de décision [171]. L'implication d'experts d'autres spécialités telles que la psychologie, la sociologie, etc. constituerait un atout majeur pour affiner les facteurs utilisés pour l'entraînement des algorithmes. Par exemple, un psychologue sera en mesure de mieux caractériser le facteur « stress », ce qui permettra d'établir des catégories plus précises, améliorant ainsi la précision du modèle prédictif. Il serait également intéressant de développer un CDSS permettant la prédiction en intégrant les modèles créés sur la base de la nouvelle approche hybride et permettant une communication efficace médecin-algorithme. Il faudrait alors que le CDSS offre une explicabilité des décisions suggérées par le modèle de prédiction, et qu'il soit transparent en termes de données et de performances des modèles employés. Plusieurs méthodes d'explicabilité et d'interprétation des algorithmes d'IA sont disponibles dans la littérature [172] [173] telles que, SHAP - *SHapley Additive exPlanations* [174], Anchors [175], LIME - *Local Interpretable Model-agnostic Explanations* [176]. Leur usage avec des modèles basés sur notre approche pourrait apporter plus de pertinence et de précision, étant donné que les facteurs utilisés sont catégorisés sur la base de connaissances fondées sur la pratique médicale. L'application des méthodes de recherche à une mise en œuvre clinique constitue une frontière de la recherche [177], en particulier pour rendre les résultats de l'explicabilité appréhendables par les médecins. Par conséquent, il serait utile d'évaluer avec les médecins dans quelle mesure leur implication dans la conception de systèmes basés sur l'IA peut améliorer l'explicabilité et lever les obstacles à l'utilisation de ces systèmes.

Il est courant de considérer les avancées numériques, telles que *ChatGPT*, comme la dernière révolution en matière de collaboration humain-machine. Or, *ChatGPT* illustre parfaitement le paradoxe de la collaboration humain-machine. Il est capable de fournir des réponses pertinentes à des questions posées en langage naturel, suggérant une interaction humain-machine naturelle. Cependant, il reste limité par ses capacités programmées et ne peut pas vraiment comprendre les nuances de la langue et les contextes culturels ou émotionnels. Ainsi, bien que *ChatGPT* semble collaborer étroitement avec les humains, il n'est pas réellement capable de les comprendre. Ce paradoxe met en évidence les limites des machines et souligne l'importance de la prise en compte des facteurs

humains dans la collaboration humain-machine.

L'intelligence collective est plus prometteuse pour promouvoir un avenir plus sûr de l'évolution des sciences avec le numérique. La collaboration entre humains est capable d'améliorer la qualité des décisions grâce à la diversité des idées et des perspectives. Ces capacités humaines peuvent être augmentées, notamment pour résoudre des problèmes complexes, par la puissance de la technologie numérique. L'IA, quant à elle, est capable de traiter des quantités massives de données et d'identifier des tendances et des modèles complexes. La combinaison des capacités de traitement de l'information de l'IA et de l'expérience humaine pourra permettre de fonder des décisions plus éclairées et plus précises. Cette complémentarité des compétences et des capacités humaines et de l'IA permet d'offrir des services plus personnalisés et plus pertinents. De plus, le sentiment de participation et de collaboration est très propice à l'amélioration de l'acceptabilité humaine vis-à-vis des solutions numériques. Par ailleurs, il est primordial que ces solutions reposent sur des données fiables et restent sous le contrôle de l'humain.



# RÉFÉRENCES

- [1] Chadia Ed-Driouch, Franck Mars, Pierre-Antoine Gourraud, and Cédric Dumas. Addressing the challenges and barriers to the integration of machine learning into clinical practice : An innovative method to hybrid human-machine intelligence. *Sensors*, 22(21) :8313, 1 2022. ISSN 1424-8220. doi : 10.3390/s22218313.
- [2] Chadia Ed-Driouch, Florent Chéneau, Françoise Simon, Guillaume Pasquier, Benoit Combès, Anne Kerbrat, Emmanuelle Le Page, Sophie Limou, Nicolas Vince, David-Axel Laplaud, Franck Mars, Cédric Dumas, Gilles Edan, and Pierre-Antoine Gourraud. Multiple sclerosis clinical decision support system based on projection to reference datasets. *Annals of Clinical and Translational Neurology*, 9(12) :1863–1873, 12 2022. doi : 10.1002/acn3.51649.
- [3] Chadia Ed-Driouch, Pierre-Antoine Gourraud, Cédric Dumas, and Franck Mars. The integration of human intelligence into artificial intelligence to provide medical practice-based predictions. *HHAI 2022 : Augmenting Human Intellect*, pages 293–294, 2022. doi : 10.3233/FAIA220221.
- [4] Chadia Ed-Driouch, Gilles Edan, Franck Mars, Cédric Dumas, and Pierre-Antoine Gourraud. L’interaction homme-machine au service de la médecine de précision. 6 2022. <https://hal.science/hal-03715833>.
- [5] Chadia Ed-driouch, Franck Mars, Pierre-Antoine Gourraud, and Cédric Dumas. L’intelligence hybride pour prédire l’évolution des maladies chroniques : Hybrid intelligence to predict chronic disease progression. In *Adjunct Proceedings of the 33rd Conference on l’Interaction Humain-Machine, IHM ’22 Adjunct*, pages 1–5. ACM, 2022. ISBN 978-1-4503-9198-6. doi : 10.1145/3502178.3529109.
- [6] Mohamed H. Shahin, Sanchita Bhattacharya, Diego Silva, Sarah Kim, Jackson Burton, Jagdeep Podichetty, Klaus Romero, and Daniela J. Conrado. Open data revolution in clinical research : Opportunities and challenges. *Clinical and Translational Science*, 13(4) :665–674, 2020. ISSN 1752-8062. doi : 10.1111/cts.12756.

- [7] Chris Giordano, Meghan Brennan, Basma Mohamed, Parisa Rashidi, François Modave, and Patrick Tighe. Accessing artificial intelligence for clinical decision-making. *Frontiers in Digital Health*, 3 :645232, 6 2021. ISSN 2673-253X. doi : 10.3389/fdgth.2021.645232.
- [8] Alison Cave, Xavier Kurz, and Peter Arlett. Real-world data for regulatory decision making : Challenges and possible solutions for europe. *Clinical Pharmacology Therapeutics*, 106(1) :36–39, 2019. ISSN 1532-6535. doi : 10.1002/cpt.1426.
- [9] Sahil Sandhu, Anthony L. Lin, Nathan Brajer, Jessica Sperling, William Ratliff, Armando D. Bedoya, Suresh Balu, Cara O'Brien, and Mark P. Sendak. Integrating a machine learning system into clinical workflows : Qualitative study. *Journal of Medical Internet Research*, 22(11) :e22421, 11 2020. ISSN 1438-8871. doi : 10.2196/22421.
- [10] J. Flach, Peter Reynolds, C. Cao, and Tiffany N. Saffell. Engineering representations to support evidence-based clinical practice. 2017. doi : 10.1177/2327857917061015.
- [11] P. Wasi. Human genomics : implications for health. *The Southeast Asian Journal of Tropical Medicine and Public Health*, 28 Suppl 2 :19–24, 1997. ISSN 0125-1562.
- [12] Hassane Alami, Marie-Pierre Gagnon, Mohamed Ali Ag Ahmed, and Jean-Paul Fortin. Digital health : Cybersecurity is a value creation lever, not only a source of expenditure. *Health Policy and Technology*, 8(4) :319–321. ISSN 2211-8837.
- [13] Le règlement général sur la protection des données - rgpd | cnil. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees>, [Online; accessed 2023-01-23].
- [14] Mehrnaz Mashoufi, Haleh Ayatollahi, and Davoud Khorasani-Zavareh. A review of data quality assessment in emergency medical services. *The Open Medical Informatics Journal*, 12 :19–32, 5 2018. ISSN 1874-4311. doi : 10.2174/1874431101812010019.
- [15] Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. *Science (New York, N.Y.)*, 366(6464) :447–453, 10 2019. ISSN 1095-9203. doi : 10.1126/science.aax2342.

- [16] Emily Beers, Marci Lee Nilsen, and Jonas T. Johnson. The role of patients : Shared decision-making. *Otolaryngologic Clinics of North America*, 50(4) :689–708, 8 2017. ISSN 0030-6665. doi : 10.1016/j.otc.2017.03.006.
- [17] Cécile Petitgand, Aude Motulsky, Jean-Louis Denis, and Catherine Régis. Investigating the barriers to physician adoption of an artificial intelligence- based decision support system in emergency care : An interpretative qualitative study. *Studies in Health Technology and Informatics*, 270 :1001–1005, 6 2020. ISSN 1879-8365. doi : 10.3233/SHTI200312.
- [18] Reed T. Sutton, David Pincock, Daniel C. Baumgart, Daniel C. Sadowski, Richard N. Fedorak, and Karen I. Kroeker. An overview of clinical decision support systems : benefits, risks, and strategies for success. *npj Digital Medicine*, 3(1) :1–10, 2 2020. ISSN 2398-6352. doi : 10.1038/s41746-020-0221-y.
- [19] Joan S. Ash, Dean F. Sittig, Emily M. Campbell, Kenneth P. Guappone, and Richard H. Dykstra. Some unintended consequences of clinical decision support systems. *AMIA Annual Symposium Proceedings*, 2007 :26–30, 2007. ISSN 1942-597X.
- [20] Wilfred Bonney. *Impacts and Risks of Adopting Clinical Decision Support Systems*. In- techOpen, 9 2011. ISBN 978-953-307-258-6.
- [21] Doris McGartland Rubio, Ellie E. Schoenbaum, Linda S. Lee, David E. Schteingart, Paul R. Marantz, Karl E. Anderson, Lauren Dewey Platt, Adriana Baez, and Karin Esposito. Defining translational research : Implications for training. 85(3) :470–475, 3 2010. ISSN 1040-2446. doi : 10.1097/ACM.0b013e3181ccd618.
- [22] Caitlin Maloy. Library guides : Data resources in the health sciences : Clinical data. <https://guides.lib.uw.edu/hsl/data/findclin>, [Online ; accessed 2022-12-26].
- [23] Stanley C. Ahalt, Christopher G. Chute, Karamarie Fecho, Gustavo Glusman, Jennifer Hadlock, Casey Overby Taylor, Emily R. Pfaff, Peter N. Robinson, Harold Solbrig, Casey Ta, Nicholas Tatonetti, Chunhua Weng, and The Biomedical Data Translator Consortium. *Clinical and Translational Science*, 12(4) :329–333, 2019. ISSN 1752-8062. doi : 10.1111/cts.12638.

- [24] Eduardo Hariton and Joseph J. Locascio. Randomised controlled trials—the gold standard for effectiveness research. *BJOG : an international journal of obstetrics and gynaecology*, 125(13) :1716, 12 2018. ISSN 1470-0328. doi : 10.1111/1471-0528.15199.
- [25] Lawrence Blonde, Kamlesh Khunti, Stewart B. Harris, Casey Meizinger, and Neil S. Skolnik. Interpretation and impact of real-world clinical data for the practicing clinician. *Advances in Therapy*, 35(11) :1763–1774, 11 2018. ISSN 1865-8652. doi : 10.1007/s12325-018-0805-y.
- [26] Langreth and Waldholz. New era of personalized medicine : targeting drugs for each unique genetic profile. *The Oncologist*, 4(5) :426–427, 1999. ISSN 1549-490X.
- [27] Sean Xinghua Hu, Thomas Foster, and Ann Kieffaber. Pharmacogenomics and personalized medicine : mapping of future value creation. *BioTechniques*, 39(4S) :S1–S6, 10 2005. ISSN 0736-6205. doi : 10.2144/000112048.
- [28] Eugene H. Blackstone. Precision medicine versus evidence-based medicine. *Circulation*, 140(15) :1236–1238, 10 2019. doi : 10.1161/CIRCULATIONAHA.119.043014.
- [29] Neha Akhoon. Precision medicine : A new paradigm in therapeutics. *International Journal of Preventive Medicine*, 12 :12, 2 2021. ISSN 2008-7802. doi : 10.4103/ijpvm.IJPVM\_375\_19.
- [30] Inke R. König, Oliver Fuchs, Gesine Hansen, Erika von Mutius, and Matthias V. Kopp. What is precision medicine? *The European Respiratory Journal*, 50(4) :1700391, 10 2017. ISSN 1399-3003. doi : 10.1183/13993003.00391-2017.
- [31] Jan Trøst Jørgensen. Twenty years with personalized medicine : Past, present, and future of individualized pharmacotherapy. *The Oncologist*, 24(7) :e432–e440, 7 2019. ISSN 1083-7159. doi : 10.1634/theoncologist.2019-0054.
- [32] Adam A. Friedman, Anthony Letai, David E. Fisher, and Keith T. Flaherty. Precision medicine for cancer with next-generation functional diagnostics. *Nature Reviews Cancer*, 15(12) :747–756, 12 2015. ISSN 1474-1768. doi : 10.1038/nrc4015.

- [33] Xiaolan Deng and Yusuke Nakamura. Cancer precision medicine : From cancer screening to drug selection and personalized immunotherapy. *Trends in Pharmaceutical Sciences*, 38(1) :15–24, 1 2017. ISSN 1873-3735. doi : 10.1016/j.tips.2016.10.013.
- [34] Gideon M. Blumenthal, Elizabeth Mansfield, and Richard Pazdur. *JAMA Oncology*, 2(1) :13–14, 1 2016. ISSN 2374-2437. doi : 10.1001/jamaoncol.2015.4503.
- [35] Barbara C. Worst, Cornelis M. van Tilburg, Gnana Prakash Balasubramanian, Petra Fiesel, and al. Next-generation personalised medicine for high-risk paediatric cancer patients – the inform pilot study. *European Journal of Cancer*, 65 :91–101, 9 2016. ISSN 0959-8049. doi : 10.1016/j.ejca.2016.06.009.
- [36] Suzanne J. Forrest, Birgit Geoerger, and Katherine A. Janeway. Precision medicine in pediatric oncology. *Current Opinion in Pediatrics*, 30(1) :17–24, 2018. ISSN 1531-698X. doi : 10.1097/MOP.0000000000000570.
- [37] Atif A. Ahmed, Divya S. Vundamati, Midhat S. Farooqi, and Erin Guest. Precision medicine in pediatric cancer : Current applications and future prospects. *High-Throughput*, 7(4), 12 2018. ISSN 2571-5135. doi : 10.3390/ht7040039.
- [38] Richard Weinshilboum and Liewei Wang. Pharmacogenomics : Precision medicine and drug response. *Mayo Clinic proceedings*, 92(11) :1711–1722, 11 2017. ISSN 0025-6196. doi : 10.1016/j.mayocp.2017.09.001.
- [39] Zheng-Guo Wang, Liang Zhang, and Wen-Jun Zhao. Definition and application of precision medicine. *Chinese Journal of Traumatology*, 19(5) :249–250, 10 2016. ISSN 1008-1275. doi : 10.1016/j.cjtee.2016.04.005.
- [40] Richard W. Peck. Precision medicine is not just genomics : The right dose for every patient. *Annual Review of Pharmacology and Toxicology*, 58 :105–122, 1 2018. ISSN 1545-4304. doi : 10.1146/annurev-pharmtox-010617-052446.
- [41] Joshua C. Denny and Francis S. Collins. Precision medicine in 2030—seven ways to transform healthcare. *Cell*, 184(6) :1415–1419, 3 2021. ISSN 0092-8674. doi : 10.1016/j.cell.2021.01.015.



- [42] Sofia Georgopoulou, Louise Prothero, and David P. D’Cruz. Physician-patient communication in rheumatology : a systematic review. *Rheumatology International*, 38 (5) :763–775, 5 2018. ISSN 1437-160X. doi : 10.1007/s00296-018-4016-2.
- [43] John Heritage and Douglas Maynard. Problems and prospects in the study of physician-patient interaction : 30 years of research. *Annual Review of Sociology - ANNU REV SOCIOL*, 32 :351–374, 8 2006. doi : 10.1146/annurev.soc.32.082905.093959.
- [44] S Allen, SN Rogers, S Brown, and RV Harris. What are the underlying reasons behind socioeconomic differences in doctor-patient communication in head and neck oncology review clinics? *Health Expectations*, 24(1) :140–151, 2 2021. ISSN 1369-6513. doi : 10.1111/hex.13163.
- [45] M Bahadori, E Hasanpoor, M Yaghoubi, and E HaghGoshyie. Determinants of a high-quality consultation in medical communications : a systematic review of qualitative and quantitative evidence. *International Journal Of Human Rights In Health Care*, 13(5) :457–483, 11 2020. ISSN 2056-4902. doi : 10.1108/IJHRH-05-2019-0031.
- [46] E. Schmidt, L. Gramm, and E. Farin. [communication preferences of patients with chronic back pain in medical rehabilitation]. *Schmerz (Berlin, Germany)*, 26(1) :69–76, 2 2012. ISSN 1432-2129. doi : 10.1007/s00482-011-1105-5.
- [47] Mandy Moffat, Jennifer Cleland, Thys van der Molen, and David Price. Sub-optimal patient and physician communication in primary care consultations : its relation to severe and difficult asthma. *Primary Care Respiratory Journal : Journal of the General Practice Airways Group*, 15(3) :159–165, 6 2006. ISSN 1471-4418. doi : 10.1016/j.pcrj.2006.02.006.
- [48] N. Röttele, A.C. Schöpf-Lazzarino, S. Becker, M. Körner, M. Boeker, and M.A. Wirtz. Agreement of physician and patient ratings of communication in medical encounters : A systematic review and meta-analysis of interrater agreement. *Patient Education and Counseling*, 103(10) :1873–1882, 2020. doi : 10.1016/j.pec.2020.04.002.
- [49] LL DuBenske, SB Schrager, ME Hitchcock, AK Kane, TA Little, HE McDowell, and ES Burnside. Key elements of mammography shared decision-making : a scoping

review of the literature. *JOURNAL OF GENERAL INTERNAL MEDICINE*, 33(10) : 1805–1814, 10 2018. ISSN 0884-8734. doi : 10.1007/s11606-018-4576-6.

- [50] H.M. Olivieri. Recta ratio agibilium in a medical context : The role of virtue in the physician-patient relationship. *Philosophy, Ethics, and Humanities in Medicine*, 13(1), 2018. doi : 10.1186/s13010-018-0062-3.
- [51] B. Cobos, K. Haskard-Zolnierok, and K. Howard. White coat hypertension : Improving the patient–health care practitioner relationship. *Psychology Research and Behavior Management*, 8 :133–141, 2015. doi : 10.2147/PRBM.S61192.
- [52] A. Sherwood, J. Brinkmann, and S. Fatone. Review of benefits to practitioners of using good patient-practitioner communication. *Journal of Prosthetics and Orthotics*, 30(1) :5–12, 2018. doi : 10.1097/JPO.000000000000165.
- [53] S Hajebrahimi, A Janati, M Arab-Zozani, M Sokhanvar, E Haghgoshayie, Y Siraneh, M Bahadori, and E Hasanpoor. Medical visit time and predictors in health facilities : a mega systematic review and meta-analysis. *INTERNATIONAL JOURNAL OF HUMAN RIGHTS IN HEALTH CARE*, 12(5) :373–402, 11 2019. ISSN 2056-4902. doi : 10.1108/IJHRH-05-2019-0036.
- [54] M. Bahadori, M. Yaghoubi, E. Haghgoshiyie, M. Ghasemi, and E. Hasanpoor. Patients’ and physicians’ perspectives and experiences on the quality of medical consultations : A qualitative study. *International Journal of Evidence-Based Healthcare*, 18(2) :247–255, 2020. doi : 10.1097/XEB.0000000000000210.
- [55] G. Makoul. Essential elements of communication in medical encounters : the kalamazoo consensus statement. *Academic Medicine : Journal of the Association of American Medical Colleges*, 76(4) :390–393, 4 2001. ISSN 1040-2446. doi : 10.1097/00001888-200104000-00021.
- [56] J. Menichetti, H.C. Lie, A.V. Mellblom, E.A. Brembo, H. Eide, P. Gulbrandsen, L. Heyn, K.H. Saltveit, H. Strømme, V. Sundling, E. Turk, and L.K. Juvet. Tested communication strategies for providing information to patients in medical consultations : A scoping review and quality assessment of the literature. *Patient Education and Counseling*, 2021. doi : 10.1016/j.pec.2021.01.019.

- [57] H.F. Hedian, J.A. Greene, and T.M. Niessen. The electronic health record and the clinical examination. *Medical Clinics of North America*, 102(3) :475–483, 2018. doi : 10.1016/j.mcna.2017.12.009.
- [58] John Flach, Peter Reynolds, Libby Duryee, Bryan Young, and Jeff Graley. Digital healthcare : Moving beyond the data input/out problem toward enhancing clinical judgment. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 8(1) :57–61, 9 2019. ISSN 2327-8595. doi : 10.1177/2327857919081013.
- [59] P.J. Stappers and J.M. Flach. Visualizing cognitive systems : getting past block diagrams. volume 1, pages 821–826 vol.1, 10 2004. doi : 10.1109/ICSMC.2004.1398404.
- [60] Paul A. Barclay and Clint A. Bowers. Design for the illiterate : A scoping review of tools for improving the health literacy of electronic health resources. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 61(1) :545–549, 9 2017. ISSN 2169-5067. doi : 10.1177/1541931213601620.
- [61] D Riedl and G Schussler. The influence of doctor-patient communication on health outcomes : A systematic review. *Psychosom Med Psychother*, 63(2) :131–150, 2017. ISSN 1438-3608. doi : 10.13109/zptm.2017.63.2.131.
- [62] Laura Canals-Ruiz, Marta Comellas, and Luís Lizán. Preferences, satisfaction and decision-making processes in osteoporosis treatment : a systematic review of the literature. *Journal of Comparative Effectiveness Research*, 10(8) :629–645, 6 2021. ISSN 2042-6305. doi : 10.2217/cer-2020-0216.
- [63] Tom Janisse. The human-machine interface : Inviting your computer into your patient-clinician relationship. *The Permanente Journal*, 5(1) :2–4, 2001. ISSN 1552-5767.
- [64] Tim McEwen, John Flach, and Nancy Elder. Ecological interface for assessing cardiac disease. pages 881–888. American Society of Mechanical Engineers Digital Collection, 8 2013. doi : 10.1115/ESDA2012-82974.

- [65] A. T. M. Wasylewicz and A. M. J. W. Scheepers-Hoeks. Clinical decision support systems. In Pieter Kubben, Michel Dumontier, and Andre Dekker, editors, *Fundamentals of Clinical Data Science*. Springer, Cham (CH), 2019. ISBN 978-3-319-99712-4.
- [66] Qihong Ke, Jun Liu, Mohammed Bennamoun, Senjian An, Ferdous Sohel, and Farid Boussaid. Chapter 5 - computer vision for human-machine interaction. In Marco Leo and Giovanni Maria Farinella, editors, *Computer Vision for Assistive Healthcare*, Computer Vision and Pattern Recognition, pages 127–145. Academic Press, 1 2018. ISBN 978-0-12-813445-0.
- [67] Peng Zhang. Chapter 13 - human-machine interfaces. In Peng Zhang, editor, *Advanced Industrial Control Technology*, pages 527–555. William Andrew Publishing, Oxford, 1 2010. ISBN 978-1-4377-7807-6.
- [68] Evangelia Tsolaki, Patricia Svolos, Evanthia Kousi, Eftychia Kapsalaki, Ioannis Fezoulidis, Konstantinos Fountas, Kyriaki Theodorou, Constantine Kappas, and Ioannis Tsougos. Fast spectroscopic multiple analysis (fasma) for brain tumor classification : a clinical decision support system utilizing multi-parametric 3t mr data. *International Journal of Computer Assisted Radiology and Surgery*, 10(7) :1149–1166, 7 2015. ISSN 1861-6429. doi : 10.1007/s11548-014-1088-7.
- [69] Aron Simon Downie, Mark Hancock, Christina Abdel Shaheed, Andrew J. McClachlan, Ahmet Baki Kocaballi, Christopher M. Williams, Zoe A. Michaleff, and Chris G. Maher. An electronic clinical decision support system for the management of low back pain in community pharmacy : Development and mixed methods feasibility study. *JMIR medical informatics*, 8(5) :e17203, 5 2020. ISSN 2291-9694. doi : 10.2196/17203.
- [70] Guangyao Wu, Pei Yang, Yuanliang Xie, Henry C. Woodruff, and al. Development of a clinical decision support system for severity risk prediction and triage of covid-19 patients at hospital admission : an international multicentre study. *The European Respiratory Journal*, 56(2) :2001104, 8 2020. ISSN 1399-3003. doi : 10.1183/13993003.01104-2020.
- [71] Saeid Eslami, Nicolette F. de Keizer, Dave A. Dongelmans, Evert de Jonge, Marcus J.

- Schultz, and Ameen Abu-Hanna. Effects of two different levels of computerized decision support on blood glucose regulation in critically ill patients. *International Journal of Medical Informatics*, 81(1) :53–60, 1 2012. ISSN 1386-5056. doi : 10.1016/j.ijmedinf.2011.10.004.
- [72] Raja Parasuraman and Dietrich H. Manzey. Complacency and bias in human use of automation : an attentional integration. *Human Factors*, 52(3) :381–410, 6 2010. ISSN 0018-7208. doi : 10.1177/0018720810376055.
- [73] Kate Goddard, Abdul Roudsari, and Jeremy C Wyatt. Automation bias : a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1) :121–127, 1 2012. ISSN 1067-5027. doi : 10.1136/amiajnl-2011-000089.
- [74] Chiang S. Jao, Daniel B. Hier, Chiang S. Jao, and Daniel B. Hier. *Clinical Decision Support Systems : An Effective Pathway to Reduce Medical Errors and Improve Patient Safety*. IntechOpen. ISBN 978-953-7619-64-0. <https://www.intechopen.com/state.item.id>.
- [75] Mark A. Musen, Yuval Shahar, and Edward H. Shortliffe. Clinical decision-support systems. In Edward H. Shortliffe and Leslie E. Perreault, editors, *Medical Informatics : Computer Applications in Health Care and Biomedicine*, Health Informatics, pages 573–609. Springer, New York, NY, 2001. ISBN 978-0-387-21721-5. doi : 10.1007/978-0-387-21721-5\_16.
- [76] Jerome A. Osheroﬀ, Jonathan M. Teich, Blackford Middleton, Elaine B. Steen, Adam Wright, and Don E. Detmer. A roadmap for national action on clinical decision support. *Journal of the American Medical Informatics Association : JAMIA*, 14(2) :141–145, 2007. ISSN 1067-5027. doi : 10.1197/jamia.M2334.
- [77] Edward H Shortliffe. *Computer-based medical consultations : MYCIN*. Elsevier, 2006.
- [78] B.G. Buchanan and E.H. Shortliffe. Rule-based expert systems : the mycin experiments of the stanford heuristic programming project. *Addison-Wesley*, 1984.
- [79] Edward H Shortliffe and Bruce G Buchanan. Computer-based consultations in clinical therapeutics : explanation and rule acquisition capabilities of the mycin system. *Computers and Biomedical Research*, 8(4) :303–318, 1975.

- [80] Randolph A Miller, Harold E Pople Jr, and James D Myers. Internist-1, an experimental computer-based diagnostic consultant for general internal medicine. *New England Journal of Medicine*, 307(8) :468–476, 1982.
- [81] E. Pesonen, C. Ohmann, M. Eskelinen, and M. Juhola. Diagnosis of acute appendicitis in two databases. evaluation of different neighborhoods with an LVQ neural network. 37(1) :59–63. ISSN 0026-1270.
- [82] Mario Greco, Mario Stefanelli, Barbara Bazzanella, Aldo Barone, Federico Murgia, Mauro Sarti, and Mauro Torchio. Expert systems in clinical oncology : the oncocin experience. *Computers and Biomedical Research*, 18(6) :509–532, 1985.
- [83] Euan A Ashley. Toward precision medicine. *New England Journal of Medicine*, 375(1) :1–3, 2015.
- [84] Zhiyun Zhou, Andriy Fedorov, Rebecca Remis, Patrick Grossmann, Steve Pieper, Ron Kikinis, and Polina Golland. 3d deep learning for healthcare applications : A survey. *IEEE transactions on medical imaging*, 37(9) :1943–1960, 2018.
- [85] Wojciech Walach, Maria Magnusson, Winfried Schlee, and Damian Nowak. Artificial intelligence against covid-19 : an early review. *AI SOCIETY*, 35(4) :761–777, 2020.
- [86] Pavel Hamet and Johanne Tremblay. Artificial intelligence in medicine. *Metabolism : Clinical and Experimental*, 69S :S36–S40, 4 2017. ISSN 1532-8600. doi : 10.1016/j.metabol.2017.01.011.
- [87] I. Glenn Cohen, Ruben Amarasingham, Anand Shah, Bin Xie, and Bernard Lo. The legal and ethical concerns that arise from using complex predictive analytics in health care. *Health Affairs (Project Hope)*, 33(7) :1139–1147, 7 2014. ISSN 1544-5208. doi : 10.1377/hlthaff.2014.0048.
- [88] Stefano Triberti, Ilaria Durosini, and Gabriella Pravettoni. A "third wheel" effect in health decision making involving artificial entities : A psychological perspective. *Frontiers in Public Health*, 8 :117, 2020. ISSN 2296-2565. doi : 10.3389/fpubh.2020.00117.

- [89] Amit V. Khera, Mark Chaffin, Krishna G. Aragam, Mary E. Haas, Carolina Roselli, Seung Hoan Choi, Pradeep Natarajan, Eric S. Lander, Steven A. Lubitz, Patrick T. Ellinor, and Sekar Kathiresan. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature genetics*, 50(9) :1219–1224, 9 2018. ISSN 1061-4036. doi : 10.1038/s41588-018-0183-z.
- [90] Nicholas J. Schork. Artificial intelligence and personalized medicine. *Cancer treatment and research*, 178 :265–283, 2019. ISSN 0927-3042. doi : 10.1007/978-3-030-16391-4\_11.
- [91] Hassane Alami, Pascale Lehoux, Jean-Louis Denis, Aude Motulsky, Cecile Petitgand, Mathilde Savoldelli, Ronan Rouquet, Marie-Pierre Gagnon, Denis Roy, and Jean-Paul Fortin. Organizational readiness for artificial intelligence in health care : insights for decision-making and practice. *Journal of Health Organization and Management*, ahead-of-print(ahead-of-print), 12 2020. ISSN 1758-7247. doi : 10.1108/JHOM-03-2020-0074.
- [92] Jean-Christophe Bélisle-Pipon, Vincent Couture, Marie-Christine Roy, Isabelle Ganache, Mireille Goetghebeur, and I. Glenn Cohen. What makes artificial intelligence exceptional in health technology assessment? *Frontiers in Artificial Intelligence*, 4 : 736697, 2021. ISSN 2624-8212. doi : 10.3389/frai.2021.736697.
- [93] Home | think tank | european parliament. [www.europarl.europa.eu/thinktank](http://www.europarl.europa.eu/thinktank), [Online ; accessed 2022-11-25].
- [94] Yifan Xiang, Lanqin Zhao, Zhenzhen Liu, Xiaohang Wu, Jingjing Chen, Erping Long, Duoru Lin, Yi Zhu, Chuan Chen, Zhuoling Lin, and Haotian Lin. Implementation of artificial intelligence in medicine : Status analysis and development suggestions. *Artificial Intelligence in Medicine*, 102 :101780, 1 2020. ISSN 0933-3657. doi : 10.1016/j.artmed.2019.101780.
- [95] Dominik Dellermann, Philipp Ebel, Matthias Söllner, and Jan Marco Leimeister. Hybrid intelligence. *Business Information Systems Engineering*, 61(5) :637–643, 2019.
- [96] Fei-Yue Wang, Jianbo Guo, Guangquan Bu, and Jun Jason Zhang. Mutually trustworthy human-machine knowledge automation and hybrid augmented intelli-

gence : mechanisms and applications of cognition, management, and control for complex systems. *Frontiers of Information Technology Electronic Engineering*, 23(8) : 1142–1157, 8 2022. ISSN 2095-9184. doi : 10.1631/FITEE.2100418.

- [97] Lu Hong, PJ Lamberson, and Scott E Page. Hybrid predictive ensembles : Synergies between human and computational forecasts. *Journal of Social Computing*, 2(2) :89–102, 6 2021. ISSN 2688-5255. doi : 10.23919/JSC.2021.0009.
- [98] Nan-ning Zheng, Zi-yi Liu, Peng-ju Ren, Yong-qiang Ma, Shi-tao Chen, Si-yu Yu, Jian-ru Xue, Ba-dong Chen, and Fei-yue Wang. Hybrid-augmented intelligence : collaboration and cognition. *Frontiers of Information Technology Electronic Engineering*, 18(2) :153–179, 2 2017. ISSN 2095-9230. doi : 10.1631/FITEE.1700053.
- [99] James Shaw, Frank Rudzicz, Trevor Jamieson, and Avi Goldfarb. Artificial intelligence and the implementation challenge. *Journal of Medical Internet Research*, 21(7) : e13659, 7 2019. ISSN 1439-4456. doi : 10.2196/13659.
- [100] A. N. Ramesh, C. Kambhampati, J. R. T. Monson, and P. J. Drew. Artificial intelligence in medicine. *Annals of the Royal College of Surgeons of England*, 86(5) :334–338, 9 2004. ISSN 0035-8843. doi : 10.1308/147870804290.
- [101] Andre Kushniruk, Helen Monkman, Elizabeth Borycki, and Joseph Kannry. User-centered design and evaluation of clinical information systems : A usability engineering perspective. In Vimla L. Patel, Thomas G. Kannampallil, and David R. Kaufman, editors, *Cognitive Informatics for Biomedicine : Human Computer Interaction in Healthcare*, Health Informatics, pages 141–161. Springer International Publishing, Cham, 2015. ISBN 978-3-319-17272-9.
- [102] Ibrahim Habli, Tom Lawton, and Zoe Porter. Artificial intelligence in health care : accountability and safety. *Bulletin of the World Health Organization*, 98(4) :251–256, 4 2020. ISSN 1564-0604. doi : 10.2471/BLT.19.237487.
- [103] M. Clanet. Jean-martin charcot. 1825 to 1893. *International MS journal*, 15(2) :59–61, 6 2008. ISSN 1352-8963.



- [104] Birgit Reipert. Multiple sclerosis : a short review of the disease and its differences between men and women. *The Journal of Men's Health and Gender*, 1(4) :334–340, 12 2004. ISSN 1571-8913. doi : 10.1016/j.jmhg.2004.10.005.
- [105] Clare Walton, Rachel King, Lindsay Rechtman, Wendy Kaye, and al. Rising prevalence of multiple sclerosis worldwide : Insights from the atlas of ms, third edition. *Multiple Sclerosis Journal*, 26(14) :1816–1821, 12 2020. ISSN 1352-4585. doi : 10.1177/1352458520970841.
- [106] Stéphanie Foulon, Géric Maura, Marie Dalichampt, François Alla, Marc Debouverie, Thibault Moreau, and Alain Weill. Prevalence and mortality of patients with multiple sclerosis in france in 2012 : a study based on french health insurance data. *Journal of Neurology*, 264(6) :1185–1192, 2017. ISSN 0340-5354. doi : 10.1007/s00415-017-8513-0.
- [107] Chris Cotsapas, Mitja Mitrovic, and David Hafler. Multiple sclerosis. *Handbook of Clinical Neurology*, 148 :723–730, 2018. ISSN 0072-9752. doi : 10.1016/B978-0-444-64076-5.00046-6.
- [108] Daniel S. Reich, Claudia F. Lucchinetti, and Peter A. Calabresi. Multiple sclerosis. *The New England journal of medicine*, 378(2) :169–180, 1 2018. ISSN 0028-4793. doi : 10.1056/NEJMr1401483.
- [109] Bassem I. Yamout and Raed Alroughani. Multiple sclerosis. *Seminars in Neurology*, 38(2) :212–225, 4 2018. ISSN 1098-9021. doi : 10.1055/s-0038-1649502.
- [110] K. L. Munger, S. M. Zhang, E. O'Reilly, M. A. Hernán, M. J. Olek, W. C. Willett, and A. Ascherio. Vitamin d intake and incidence of multiple sclerosis. *Neurology*, 62(1) :60–65, 1 2004. ISSN 0028-3878, 1526-632X. doi : 10.1212/01.WNL.0000101723.79681.38.
- [111] Samantha S. Soldan and Paul M. Lieberman. Epstein–barr virus and multiple sclerosis. 21(1) :51–64. ISSN 1740-1534. doi : 10.1038/s41579-022-00770-5.
- [112] Lynn I. Levin, Kassandra L. Munger, Eilis J. O'Reilly, Kerstin I. Falk, and Alberto Ascherio. Primary infection with the epstein-barr virus and risk of mul-

multiple sclerosis. *Annals of Neurology*, 67(6) :824–830, 6 2010. ISSN 1531-8249. doi : 10.1002/ana.21978.

- [113] Adam E. Handel, Alexander J. Williamson, Giulio Disanto, Lahiru Handunnetthi, Gavin Giovannoni, and Sreeram V. Ramagopalan. An updated meta-analysis of risk of multiple sclerosis following infectious mononucleosis. *PloS One*, 5(9) :e12496, 9 2010. ISSN 1932-6203. doi : 10.1371/journal.pone.0012496.
- [114] Adam E. Handel, Alexander J. Williamson, Giulio Disanto, Ruth Dobson, Gavin Giovannoni, and Sreeram V. Ramagopalan. Smoking and multiple sclerosis : An updated meta-analysis. *PLOS ONE*, 6(1) :e16149, 1 2011. ISSN 1932-6203. doi : 10.1371/journal.pone.0016149.
- [115] Milena A. Gianfrancesco, Brigid Acuna, Ling Shen, Farren B.S. Briggs, Hong Quach, Kalliope H. Bellesis, Allan Bernstein, Anna K. Hedstrom, Ingrid Kockum, Lars Alfredsson, Tomas Olsson, Catherine Schaefer, and Lisa F. Barcellos. Obesity during childhood and adolescence increases susceptibility to multiple sclerosis after accounting for established genetic and environmental risk factors. *Obesity research clinical practice*, 8(5) :e435–e447, 2014. ISSN 1871-403X. doi : 10.1016/j.orcp.2014.01.002.
- [116] Maria Pia Amato, Tobias Derfuss, Bernard Hemmer, Roland Liblau, Xavier Montalban, Per Soelberg Sørensen, David H. Miller, and 2016ECTRIMS Focused Workshop Group. Environmental modifiable risk factors for multiple sclerosis : Report from the 2016 ectrims focused workshop. *Multiple Sclerosis (Houndmills, Basingstoke, England)*, 24(5) :590–603, 4 2018. ISSN 1477-0970. doi : 10.1177/1352458516686847.
- [117] Jiwon Oh, Angela Vidal-Jordana, and Xavier Montalban. Multiple sclerosis : clinical aspects. *Current Opinion in Neurology*, 31(6) :752–759, 12 2018. ISSN 1473-6551. doi : 10.1097/WCO.0000000000000622.
- [118] F. D. Lublin and S. C. Reingold. Defining the clinical course of multiple sclerosis : results of an international survey. national multiple sclerosis society (usa) advisory committee on clinical trials of new agents in multiple sclerosis. *Neurology*, 46(4) : 907–911, 4 1996. ISSN 0028-3878. doi : 10.1212/wnl.46.4.907.

- [119] Fred D. Lublin, Stephen C. Reingold, Jeffrey A. Cohen, Gary R. Cutter, and al. Defining the clinical course of multiple sclerosis : the 2013 revisions. *Neurology*, 83(3) : 278–286, 7 2014. ISSN 1526-632X. doi : 10.1212/WNL.0000000000000560.
- [120] Sylvia Klineova and Fred D. Lublin. Clinical course of multiple sclerosis. *Cold Spring Harbor Perspectives in Medicine*, 8(9) :a028928, 9 2018. ISSN 2157-1422. doi : 10.1101/cshperspect.a028928.
- [121] *Le point sur la sclérose en plaques*. Fondation ARSEP, edimark edition, 2020. ISBN 978-2-915968-58-3. <https://www.arsep.org/library/media/other/Publications/Livre-ARSEP-Le-point-sur-la-SEP-ed-2020.pdf>.
- [122] Stephen L. Stuckey, Tony D. Goh, Theresa Heffernan, and David Rowan. Hyperintensity in the subarachnoid space on flair mri. *American Journal of Roentgenology*, 189(4) :913–921, 10 2007. ISSN 0361-803X. doi : 10.2214/AJR.07.2424.
- [123] Alan J. Thompson, Brenda L. Banwell, Frederik Barkhof, William M. Carroll, and al. Diagnosis of multiple sclerosis : 2017 revisions of the mcdonald criteria. *The Lancet. Neurology*, 17(2) :162–173, 2 2018. ISSN 1474-4465. doi : 10.1016/S1474-4422(17)30470-2.
- [124] J. F. Kurtzke. Rating neurologic impairment in multiple sclerosis : an expanded disability status scale (edss). *Neurology*, 33(11) :1444–1452, 11 1983. ISSN 0028-3878. doi : 10.1212/wnl.33.11.1444.
- [125] Emmanuelle Leray, Jacqueline Yaouanq, Emmanuelle Le Page, Marc Coustans, David Laplaud, Joël Oger, and Gilles Edan. Evidence for a two-stage disability progression in multiple sclerosis. *Brain : A Journal of Neurology*, 133(Pt 7) :1900–1913, 7 2010. ISSN 1460-2156. doi : 10.1093/brain/awq076.
- [126] Benjamin L Conway, Burcu Zeydan, Ufuk Uygunoğlu, Michaela Novotna, Aksel Siva, Sean J Pittock, Elizabeth J Atkinson, Moses Rodriguez, and Orhun H Kantarci. Age is a critical determinant in recovery from multiple sclerosis relapses. *Multiple Sclerosis Journal*, 25(13) :1754–1763, 2019. doi : 10.1177/1352458518800815.

- [127] Derrick Robertson and Natalie Moreo. Disease-modifying therapies in multiple sclerosis : Overview and treatment considerations. *Federal practitioner : for the health care professionals of the VA, DoD, and PHS*, 33(6) :28–34, 2016.
- [128] R. Dobson and G. Giovannoni. Multiple sclerosis - a review. *European Journal of Neurology*, 26(1) :27–40, 1 2019. ISSN 1468-1331. doi : 10.1111/ene.13819.
- [129] Poul Jennum, Benedikte Wanscher, Jette Frederiksen, and Jakob Kjellberg. The socioeconomic consequences of multiple sclerosis : A controlled national study. *European Neuropsychopharmacology*, 22(1) :36–43, 1 2012. ISSN 0924-977X. doi : 10.1016/j.euroneuro.2011.05.001.
- [130] A Chan and CS Heck. Mobility in multiple sclerosis : More than just a physical problem. *International Journal of MS Care*, 2(1) :51–61, 3 2000. ISSN 1537-2073. doi : 10.7224/1537-2073-2.1.51.
- [131] Maurizio Pompili, Alberto Forte, Mario Palermo, Henry Stefani, Dorian A. Lamis, Gianluca Serafini, Mario Amore, and Paolo Girardi. Suicide risk in multiple sclerosis : a systematic review of current literature. *Journal of Psychosomatic Research*, 73 (6) :411–417, 12 2012. ISSN 1879-1360. doi : 10.1016/j.jpsychores.2012.09.011.
- [132] Aida Hougaard Andersen, Elisabeth Assing Hvidt, Niels Christian Hvidt, and Kirsten K. Roessler. ‘maybe we are losing sight of the human dimension’ – physicians’ approaches to existential, spiritual, and religious needs among patients with chronic pain or multiple sclerosis. a qualitative interview-study. *Health Psychology and Behavioral Medicine*, 8(1) :248–269, 1 2020. doi : 10.1080/21642850.2020.1792308.
- [133] Jessie Dezutter, Martin Offenbaecher, Miguel A. Vallejo, Siebrecht Vanhooren, Evalyne Thauvoye, and Loren Toussaint. Chronic pain care. *International Journal of Psychiatry in Medicine*, 51(6) :563–575, 8 2016. ISSN 1541-3527. doi : 10.1177/0091217417696738.
- [134] Heidrun Golla, Maren Galushko, Holger Pfaff, and Raymond Voltz. Unmet needs of severely affected multiple sclerosis patients : The health professionals’ view. *Palliative Medicine*, 26(2) :139–151, 3 2012. ISSN 0269-2163. doi : 10.1177/0269216311401465.

- [135] Rayan Alshamrani, Ashrf Althbiti, Yara Alshamrani, Fatimah Alkomah, and Xiaogang Ma. Model-driven decision making in multiple sclerosis research : Existing works and latest trends. *Patterns*, 1(8) :100121, 11 2020. ISSN 2666-3899. doi : 10.1016/j.patter.2020.100121.
- [136] Mário Veloso. An agent-based simulation model for informed shared decision making in multiple sclerosis. *Multiple Sclerosis and Related Disorders*, 2(4) :377–384, 10 2013. ISSN 2211-0348. doi : 10.1016/j.msard.2013.04.001.
- [137] Pierre-Antoine Gourraud, Roland Henry, Bruce AC Cree, Jason C Crane, and al. Precision medicine in chronic disease management : the ms bioscreen. *Annals of neurology*, 76(5) :633–642, 11 2014. ISSN 0364-5134. doi : 10.1002/ana.24282.
- [138] Erica Schleimer, Jennifer Pearce, Andrew Barnecut, William Rowles, Antoine Lizee, Arno Klein, Valerie J. Block, Adam Santaniello, Adam Renschen, Refujia Gomez, Anisha Keshavan, Jeffrey M. Gelfand, Roland G. Henry, Stephen L. Hauser, and Riley Bove. A precision medicine tool for patients with multiple sclerosis (the open MS BioScreen) : Human-centered design and development. 22(7) :e15605. ISSN 1438-8871. doi : 10.2196/15605.
- [139] Peter A. Calabresi, Bernd C. Kieseier, Douglas L. Arnold, Laura J. Balcer, Alexey Boyko, Jean Pelletier, Shifang Liu, Ying Zhu, Ali Seddighzadeh, Serena Hung, Aaron Deykin, and ADVANCE Study Investigators. Pegylated interferon -1a for relapsing-remitting multiple sclerosis (advance) : a randomised, phase 3, double-blind study. *The Lancet. Neurology*, 13(7) :657–665, 7 2014. ISSN 1474-4465. doi : 10.1016/S1474-4422(14)70068-7.
- [140] OFSEP - cohorte HD, . URL <https://www.ofsep.org/fr/cohorte-hd>.
- [141] Caspar E. P. van Munster and Bernard M. J. Uitdehaag. Outcome measures in clinical trials for multiple sclerosis. *CNS Drugs*, 31(3) :217–236, 2017. ISSN 1172-7047. doi : 10.1007/s40263-017-0412-5.
- [142] Tatjana Reynders, Miguel D’haeseleer, Jacques De Keyser, Guy Nagels, and Marie B. D’hooghe. Definition, prevalence and predictive factors of benign multiple

- sclerosis. *eNeurologicalSci*, 7 :37–43, 6 2017. ISSN 2405-6502. doi : 10.1016/j.ensci.2017.05.002.
- [143] E. Leray, T. Moreau, A. Fromont, and G. Edan. Epidemiology of multiple sclerosis. *Revue Neurologique*, 172(1) :3–13, 1 2016. ISSN 0035-3787. doi : 10.1016/j.neurol.2015.10.006.
- [144] Tomas Kalincik. Multiple sclerosis relapses : Epidemiology, outcomes and management. a systematic review. *Neuroepidemiology*, 44(4) :199–214, 2015. ISSN 0251-5350, 1423-0208. doi : 10.1159/000382130.
- [145] Nicola De Stefano, Maria Stromillo, Antonio Giorgio, Marco Battaglini, Maria Bartolozzi, Maria Amato, and Maria Pia Sormani. Long-term assessment of no evidence of disease activity in relapsing-remitting ms. *Neurology*, 85, 10 2015. doi : 10.1212/WNL.0000000000002105.
- [146] Maria Pia Sormani, Claudio Gasperini, Marzia Romeo, Jordi Rio, and al. Assessing response to interferon- in a multicenter dataset of patients with ms. *Neurology*, 87 (2) :134–140, 7 2016. ISSN 1526-632X. doi : 10.1212/WNL.0000000000002830.
- [147] Vittorio Martinelli, M Rodegher, Lucia Moiola, and Giacomo Comi. Late onset multiple sclerosis : Clinical characteristics, prognostic factors and differential diagnosis. *Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, 25 Suppl 4 :S350–5, 12 2004. doi : 10.1007/s10072-004-0339-8.
- [148] Saul J Weiner. Contextualizing medical decisions to individualize care. *Journal of General Internal Medicine*, 19(3) :281–285, 3 2004. ISSN 0884-8734. doi : 10.1111/j.1525-1497.2004.30261.x.
- [149] Jeremy M. G. Taylor. A comparison between a parametric and a non-parametric method of estimating the percentiles of a difference. *Biometrical Journal*, 30(5) :553–559, 1988. ISSN 1521-4036. doi : 10.1002/bimj.4710300507.
- [150] Sabine K. Schnabel and Paul H. C. Eilers. Simultaneous estimation of quantile curves using quantile sheets. *ASTA Advances in Statistical Analysis*, 97(1) :77–87, 1 2013. ISSN 1863-818X. doi : 10.1007/s10182-012-0198-1.

- [151] D. Mikis Stasinopoulos, Robert Rigby, Gillian Heller, Vlasios Voudouris, and Fernanda De Bastiani. *Flexible regression and smoothing : using GAMLSS in R*. Chapman and Hall/CRC the R Series. Chapman Hall/CRC, Boca Raton, Florida, 2017. ISBN 978-1-138-19790-9.
- [152] Getting started with quantile regression | university of virginia library research data services + sciences. <https://data.library.virginia.edu/getting-started-with-quantile-regression/>, [Online; accessed 2022-11-07].
- [153] Carine Lallemand and Guillaume Gronier. *Méthodes de design UX*. Collection Design web. EYROLLES, septembre 2018.
- [154] C. Lallemand, V. Koenig, G. Gronier, and R. Martin. Création et validation d’une version française du questionnaire attrakdiff pour l’évaluation de l’expérience utilisateur des systèmes interactifs. *European Review of Applied Psychology*, 65(5) :239–252, 9 2015. ISSN 1162-9088. doi : 10.1016/j.erap.2015.08.002.
- [155] OFSEP - RHU PRIMUS, . URL <https://www.ofsep.org/fr/rhu-primus>.
- [156] Douglas C. Engelbart and William K. English. A research center for augmenting human intellect. AFIPS ’68 (Fall, part I), page 395–410, New York, NY, USA, 1968. Association for Computing Machinery. ISBN 978-1-4503-7899-4. doi : 10.1145/1476589.1476645.
- [157] Stijn Denissen, Oliver Y. Chén, Johan De Mey, Maarten De Vos, Jeroen Van Schependom, Diana Maria Sima, and Guy Nagels. Towards multimodal machine learning prediction of individual cognitive evolution in multiple sclerosis. *Journal of Personalized Medicine*, 11(12) :1349, 12 2021. ISSN 2075-4426. doi : 10.3390/jpm11121349.
- [158] Luke Hone, Gavin Giovannoni, Ruth Dobson, and Benjamin Meir Jacobs. Predicting multiple sclerosis : Challenges and opportunities. *Frontiers in Neurology*, 12, 2022. ISSN 1664-2295.
- [159] Andrea Tacchella, Silvia Romano, Michela Ferraldeschi, Marco Salvetti, Andrea Zaccaria, Andrea Crisanti, and Francesca Grassi. Collaboration between a human group and artificial intelligence can improve prediction of multiple sclerosis

course : a proof-of-principle study. *F1000Research*, 6 :2172, 8 2018. ISSN 2046-1402. doi : 10.12688/f1000research.13114.2.

- [160] Gianluca Demartini. Hybrid human–machine information systems : Challenges and opportunities. *Computer Networks*, 90 :5–13, 10 2015. ISSN 1389-1286. doi : 10.1016/j.comnet.2015.05.018.
- [161] Abhimanyu S. Ahuja. The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ*, 7, 10 2019. ISSN 2167-8359. doi : 10.7717/peerj.7702.
- [162] RH Gross, SH Sillau, AE Miller, C Farrell, and SC Krieger. The multiple sclerosis severity score : fluctuations and prognostic ability in a longitudinal cohort of patients with ms. *Multiple Sclerosis Journal - Experimental, Translational and Clinical*, 5 (1), 3 2019. ISSN 2055-2173. doi : 10.1177/2055217319837254.
- [163] Loren A. Rolak. Multiple sclerosis : It’s not the disease you thought it was. *Clinical Medicine and Research*, 1(1) :57–60, 1 2003. ISSN 1539-4182.
- [164] J E T Akinsola. Supervised machine learning algorithms : Classification and comparison. *International Journal of Computer Trends and Technology (IJCTT)*, 48 :128–138, 6 2017. doi : 10.14445/22312803/IJCTT-V48P126.
- [165] Omer Sagi and Lior Rokach. Ensemble learning : A survey. *WIREs Data Mining and Knowledge Discovery*, 8(4) :e1249, 2018. ISSN 1942-4795. doi : 10.1002/widm.1249.
- [166] Jiawen Kong, Wojtek Kowalczyk, Duc Anh Nguyen, Thomas Bäck, and Stefan Menzel. Hyperparameter optimisation for improving classification under class imbalance. pages 3072–3078, 12 2019. doi : 10.1109/SSCI44817.2019.9002679.
- [167] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2), 2012.
- [168] Corinna Cortes and Mehryar Mohri. Confidence intervals for the area under the roc curve. volume 17. MIT Press, 2004.
- [169] Krzysztof J. Cios and G. William Moore. Uniqueness of medical data mining. *Artificial Intelligence in Medicine*, 26(1) :1–24, 9 2002. ISSN 0933-3657. doi : 10.1016/S0933-3657(02)00049-0.



- [170] Ariadne A. Nichol, Meghan C. Halley, Carole A. Federico, Mildred K. Cho, and Pamela L. Sankar. Not in my ai : Moral engagement and disengagement in health care ai development. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 28 :496–506, 2023. ISSN 2335-6928.
- [171] Max Wolf, Jens Krause, Patricia A. Carney, Andy Bogart, and Ralf H. J. M. Kurvers. Collective intelligence meets medical decision-making : The collective outperforms the best radiologist. *PLOS ONE*, 10(8) :e0134269, 2015. ISSN 1932-6203. doi : 10.1371/journal.pone.0134269.
- [172] Serg Masís. *Interpretable Machine Learning with Python : Learn to Build Interpretable High-performance Models with Hands-on Real-world Examples*. Packt Publishing, 2021. ISBN 978-1-80020-390-7.
- [173] Christoph Molnar. *Interpretable Machine Learning*. <https://christophm.github.io/interpretable-ml-book/>, [Online; accessed 2023-02-02].
- [174] Scott Lundberg and Su-In Lee. A unified approach to interpreting model predictions. 11 2017. doi : 10.48550/arXiv.1705.07874. URL <http://arxiv.org/abs/1705.07874>.
- [175] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. Anchors : High-precision model-agnostic explanations. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), 4 2018. ISSN 2374-3468. doi : 10.1609/aaai.v32i1.11491.
- [176] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?" : Explaining the predictions of any classifier. 8 2016. doi : 10.48550/arXiv.1602.04938.
- [177] Christopher J. Kelly, Alan Karthikesalingam, Mustafa Suleyman, Greg Corrado, and Dominic King. Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17(1) :195, 10 2019. ISSN 1741-7015. doi : 10.1186/s12916-019-1426-2.
- [178] Jerome Friedman Trevor Hastie, Robert Tibshirani. *Elements of statistical learning : data mining, inference, and prediction*. 2nd edition. <https://hastie.su.domains/ElemStatLearn/>, [Online; accessed 2023-02-02].

- [179] Trevor Hastie Rob Tibshirani Gareth James, Daniela Witten. An introduction to statistical learning. <https://www.statlearning.com>, [Online; accessed 2023-02-02].
- [180] C. Zhang and Y. Ma. *Ensemble machine learning : Methods and applications*. page : 329, DOI : 10.1007/9781441993267.



# **Annexes**

# ANNEXE A : FIGURES ET TABLEAUX ANNEXES



FIGURE A.1 – Graphe par paire montrant la façon dont les facteurs utilisés par le médecin diffèrent en fonction de l'EDSS

*La diagonale montre la distribution des variables individuelles pour chaque classe de l'EDSS.*

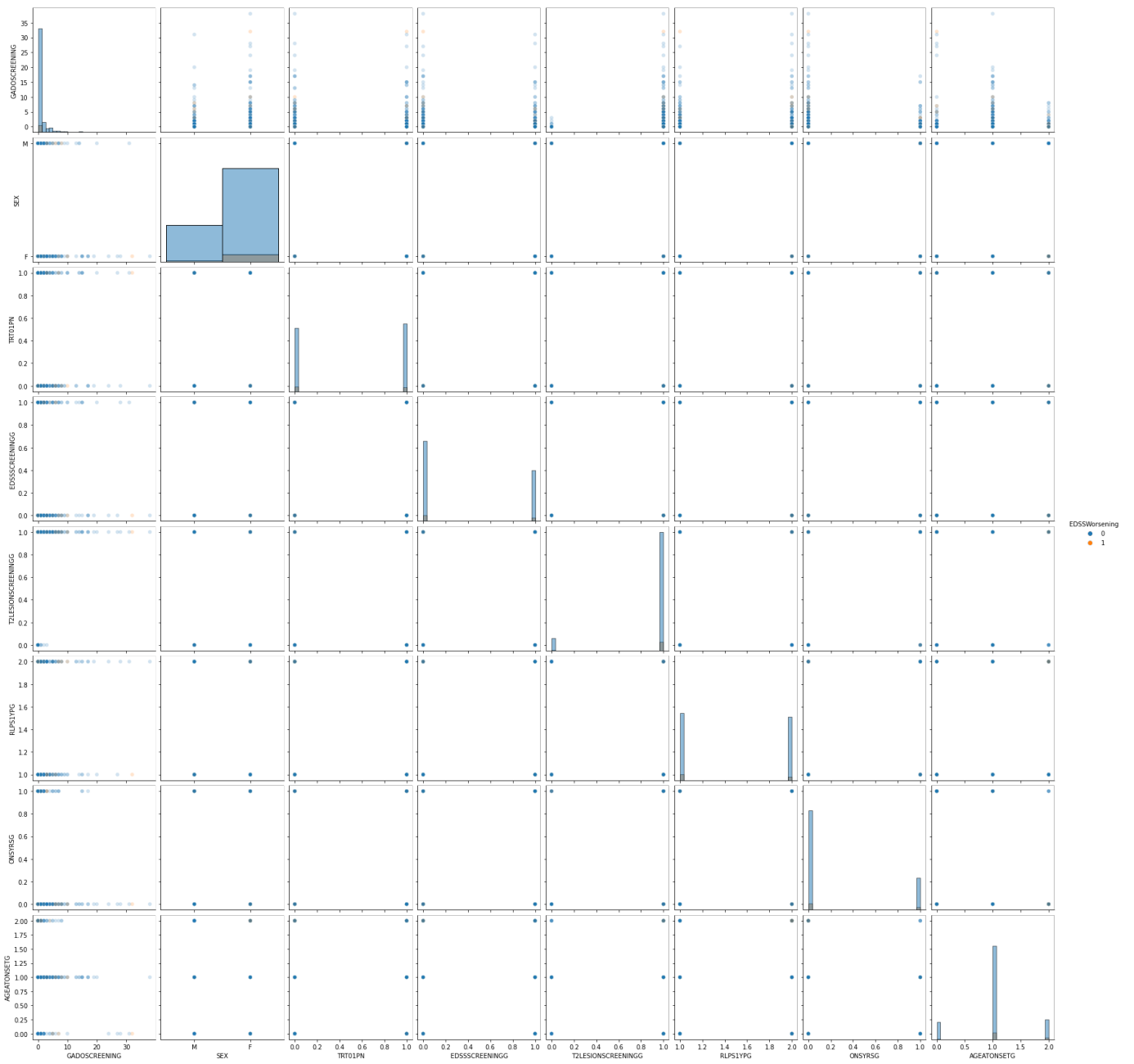


FIGURE A.2 – Graphe par paire montrant la façon dont les facteurs utilisés et catégorisés par le médecin diffèrent en fonction de l'EDSS

*La diagonale montre la distribution des variables individuelles pour chaque classe de l'EDSS.*

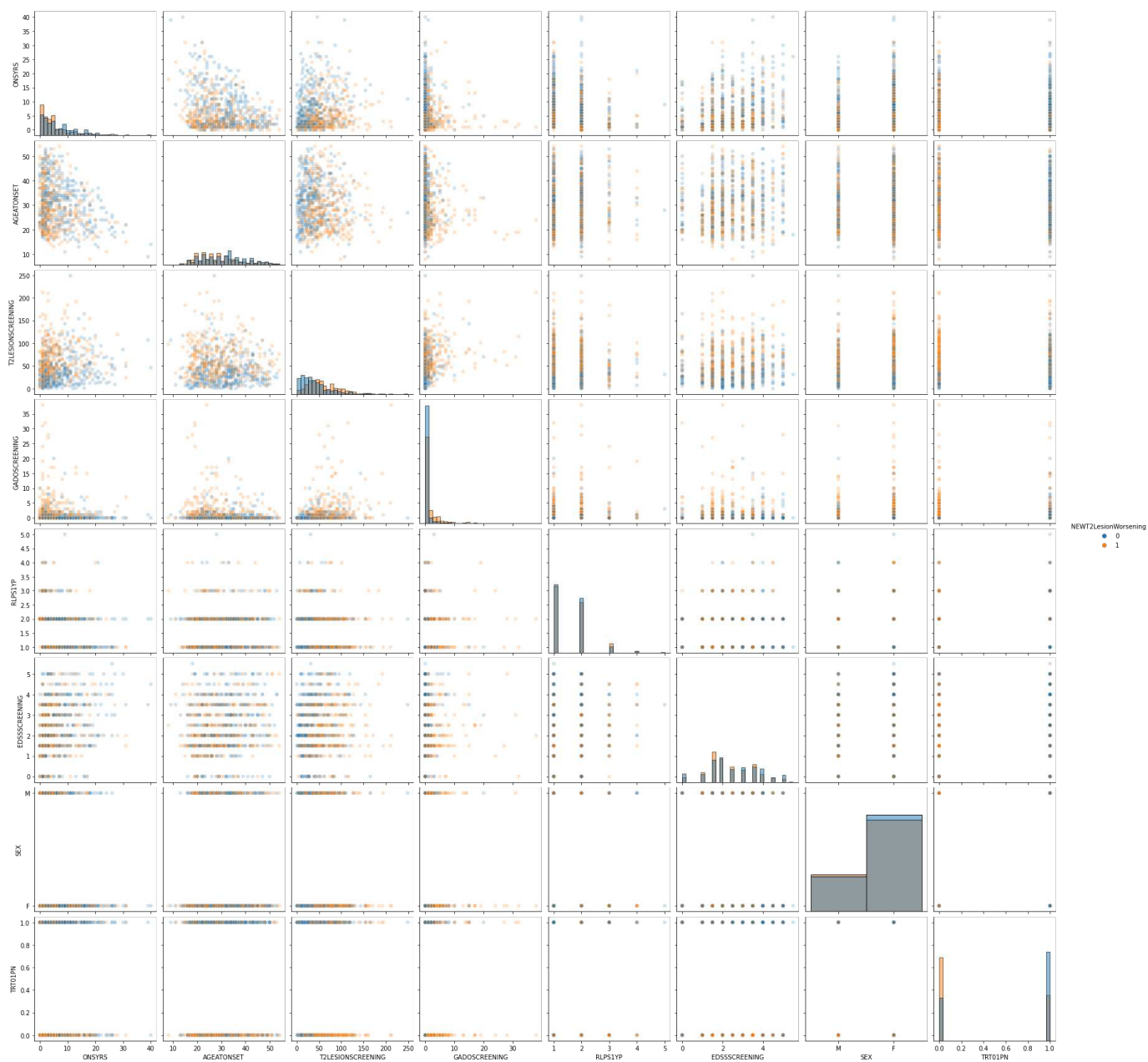


FIGURE A.3 – Graphe par paire montrant la façon dont les facteurs utilisés par le médecin diffèrent en fonction des nouvelles lésions

*La diagonale montre la distribution des variables individuelles pour chaque classe des nouvelles lésions.*

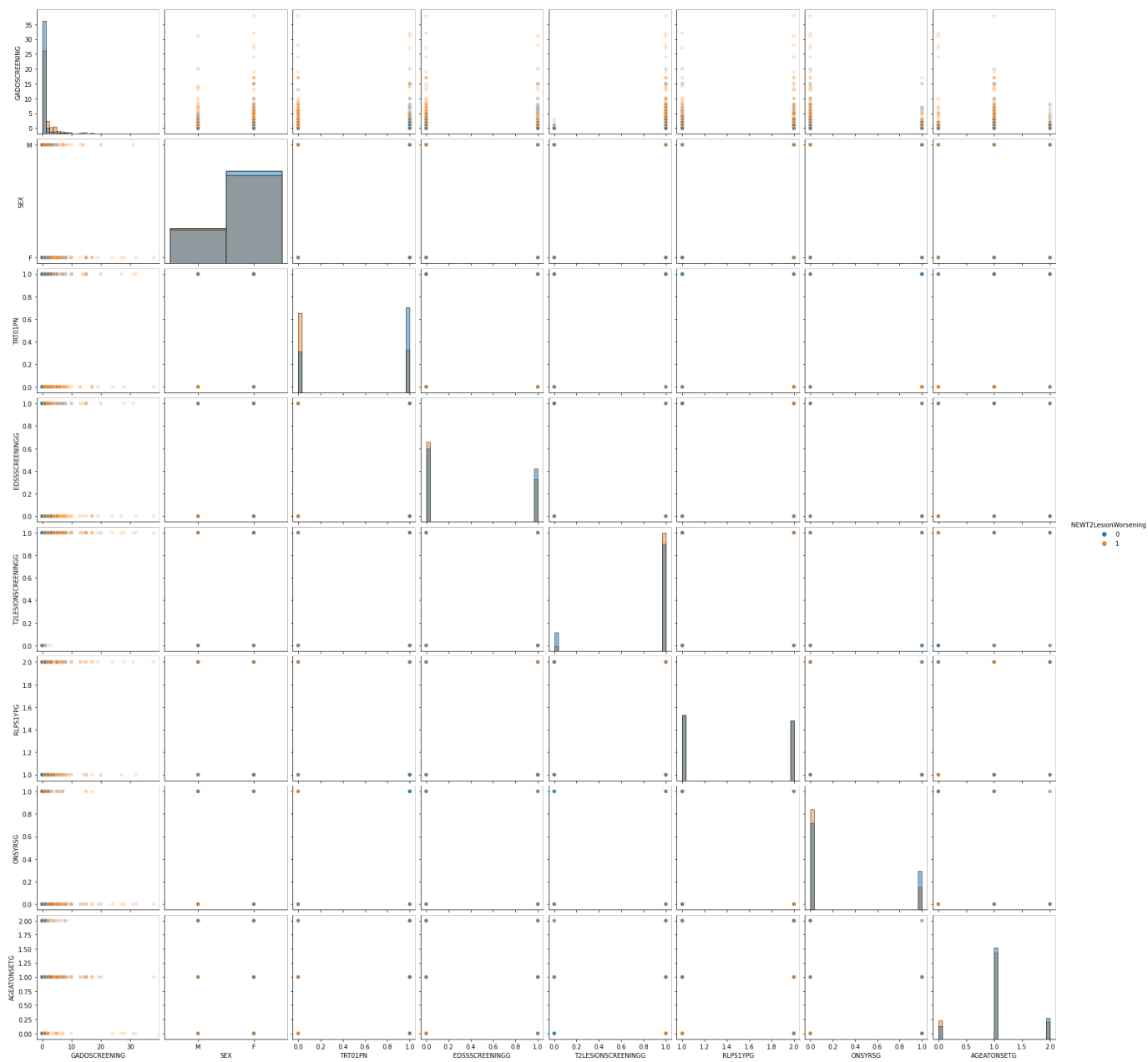


FIGURE A.4 – Graphe par paire montrant la façon dont les facteurs utilisés et catégorisés par le médecin diffèrent en fonction des nouvelles lésions.

*La diagonale montre la distribution des variables individuelles pour chaque classe des nouvelles lésions.*



TABLEAU A.1 – Statistiques descriptives des données quantitatives

|                   | <b>mean</b>  | <b>std</b>   | <b>min</b> | <b>25%</b>   | <b>50%</b>   | <b>75%</b>   | <b>max</b>  |
|-------------------|--------------|--------------|------------|--------------|--------------|--------------|-------------|
| ONSYRS            | 6.497403e+00 | 6.363143     | 0.00       | 2.000000e+00 | 4.000000e+00 | 9.000000e+00 | 40.0000     |
| AGEATONSET        | 3.011948e+01 | 9.259797     | 8.00       | 2.300000e+01 | 2.900000e+01 | 3.600000e+01 | 54.0000     |
| DIAGYRS           | 3.567532e+00 | 4.745402     | 0.00       | 1.000000e+00 | 1.000000e+00 | 5.000000e+00 | 40.0000     |
| AGEATDIAGYRS      | 3.304935e+01 | 9.513807     | 13.00      | 2.600000e+01 | 3.300000e+01 | 4.000000e+01 | 58.0000     |
| T2LESIONSCREENING | 4.965065e+01 | 36.846517    | 0.00       | 2.225000e+01 | 4.200000e+01 | 7.100000e+01 | 249.0000    |
| NormalBrainVolume | 1.580130e+06 | 96188.787187 | 1278789.06 | 1.518173e+06 | 1.586498e+06 | 1.649965e+06 | 1851174.430 |
| T2VOLUMESCREENING | 9.861367e+00 | 11.662986    | 0.00       | 1.960500e+00 | 5.734500e+00 | 1.348200e+01 | 99.054      |
| GADOSCREENING     | 1.370130e+00 | 3.652584     | 0.00       | 0.000000e+00 | 0.000000e+00 | 1.000000e+00 | 38.0000     |
| GDVOLUMESCREENING | 1.822369e-01 | 0.509063     | 0.00       | 0.000000e+00 | 0.000000e+00 | 1.230000e-01 | 5.019       |
| T1LESIONSCREENING | 2.725714e+01 | 27.619208    | 0.00       | 7.000000e+00 | 1.800000e+01 | 4.100000e+01 | 204.0000    |
| T1VOLUMESCREENING | 3.007651e+00 | 4.771851     | 0.00       | 3.157500e-01 | 1.125000e+00 | 3.732000e+00 | 39.201      |
| RLPSIYP           | 1.571429e+00 | 0.675189     | 1.00       | 1.000000e+00 | 1.000000e+00 | 2.000000e+00 | 5.0000      |
| EDSSSCREENING     | 2.442857e+00 | 1.207629     | 0.00       | 1.500000e+00 | 2.000000e+00 | 3.500000e+00 | 5.5000      |
| BOWLBLDRSCREENING | 4.415584e-01 | 0.642927     | 0.00       | 0.000000e+00 | 0.000000e+00 | 1.000000e+00 | 3.0000      |
| BRAINSTMSCREENING | 5.818182e-01 | 0.766110     | 0.00       | 0.000000e+00 | 0.000000e+00 | 1.000000e+00 | 4.0000      |
| CEREBELRSCREENING | 1.022078e+00 | 0.976062     | 0.00       | 0.000000e+00 | 1.000000e+00 | 2.000000e+00 | 4.0000      |
| CEREBRALSCREENING | 4.324675e-01 | 0.673662     | 0.00       | 0.000000e+00 | 0.000000e+00 | 1.000000e+00 | 2.0000      |

TABLEAU A.1 - Statistiques descriptives des données quantitatives (Suite)

|                   | <b>mean</b>  | <b>std</b> | <b>min</b> | <b>25%</b>   | <b>50%</b>   | <b>75%</b>   | <b>max</b> |
|-------------------|--------------|------------|------------|--------------|--------------|--------------|------------|
| DISTWALKSCREENING | 4.861039e+02 | 53.137108  | 150.00     | 5.000000e+02 | 5.000000e+02 | 5.000000e+02 | 500.000    |
| PYRAMIDSCREENING  | 1.579221e+00 | 1.031158   | 0.00       | 1.000000e+00 | 1.000000e+00 | 2.000000e+00 | 5.000      |
| SENSORYSCREENING  | 8.675325e-01 | 0.975972   | 0.00       | 0.000000e+00 | 1.000000e+00 | 2.000000e+00 | 4.000      |
| VISUALSCREENING   | 5.948052e-01 | 0.859113   | 0.00       | 0.000000e+00 | 0.000000e+00 | 1.000000e+00 | 4.000      |

TABLEAU A.2 – Description des algorithmes de ML utilisés et du rôle de certains de leurs paramètres clés [178–180]

| Algorithme                      | Description  | Paramètres clés                | Rôles des paramètres   |
|---------------------------------|--|--------------------------------|--|
| Support vector classifier (SVC) | SVC est un algorithme de classification supervisée. Il utilise des hyperplans pour séparer les données en deux classes. Le type du noyau (ex, linéaire, polynomial ou gaussienne) détermine la forme de l'hyperplan et la pénalité de régularisation régularise le modèle pour éviter l'overfitting. | kernel, C, gamma, class_weight | - <i>kernel</i> spécifie le type du noyau utilisé pour la transformation des données<br>- C est un paramètre de régularisation pour la marge d'erreur<br>- <i>gamma</i> définit la forme de la fonction utilisée pour la séparation des classes<br>- <i>class_weight</i> permet de gérer les inégalités de poids entre les différentes classes |

TABLEAU A.2 - Description des algorithmes de ML utilisés et du rôle de certains de leurs paramètres clés [178–180] (Suite)

| Algorithme                                  | Description   | Paramètres clés                  | Rôles des paramètres  |
|---|---|----------------------------------|---|
| Logistic Regression (Régression Logistique) | La régression logistique est un algorithme de classification supervisée. Il prédit la probabilité d'appartenance à une classe en utilisant une fonction logistique. La pénalité de régularisation (L1 ou L2) aide à éviter l'overfitting. | penalty, C, solver, class_weight | <ul style="list-style-type: none"> <li>- <i>penalty</i> spécifie le type de pénalité utilisée pour la régularisation L1 ou L2. L1 (pénalité Lasso) minimise la somme absolue des coefficients, tandis que L2 (pénalité Ridge) minimise la somme des coefficients au carré</li> <li>- C est un paramètre de régularisation pour la marge d'erreur</li> <li>- <i>solver</i> définit l'algorithme utilisé pour trouver les coefficients optimaux</li> <li>- <i>class_weight</i> permet de gérer les inégalités de poids entre les différentes classes</li> </ul> |

TABLEAU A.2 - Description des algorithmes de ML utilisés et du rôle de certains de leurs paramètres clés [178–180] (Suite)

| Algorithme                              | Description  | Paramètres clés   | Rôles des paramètres  |
|---|--|---|---|
| Decision Tree<br>(Arbre de<br>Décision) | Un arbre de décision est un algorithme de classification supervisé. Il divise répétitivement les données en sous-groupes plus petits en fonction de la caractéristique qui maximise la séparation des classes. La profondeur de l'arbre peut être limitée pour éviter l'overfitting. | criterion,<br>max_depth,<br>min_samples_split,<br>min_samples_leaf,<br>class_weight | <ul style="list-style-type: none"> <li>- <i>criterion</i> spécifie la mesure utilisée pour déterminer la qualité de la séparation des données (par exemple, entropie ou impureté gini)</li> <li>- <i>max_depth</i> définit la profondeur maximale de l'arbre</li> <li>- <i>min_samples_split</i> spécifie le nombre minimum d'échantillons nécessaires pour effectuer une séparation</li> <li>- <i>min_samples_leaf</i> détermine le nombre minimum d'échantillons requis pour être considérés comme une feuille de l'arbre</li> <li>- <i>class_weight</i> permet de gérer les inégalités de poids entre les différentes classes</li> </ul> |

TABLEAU A.2 - Description des algorithmes de ML utilisés et du rôle de certains de leurs paramètres clés [178–180] (Suite)

| Algorithme | Description  | Paramètres clés  | Rôles des paramètres  |
|------------|--|--|---|
| Bagging    | <p>Bagging est un algorithme d'ensemble. Il fonctionne en générant plusieurs échantillons aléatoires du jeu de données et en entraînant un modèle séparé sur chaque échantillon. Les résultats des différents modèles sont ensuite combinés pour faire des prédictions. Le but est d'améliorer la stabilité et la robustesse des prédictions en utilisant plusieurs modèles.</p> | <p>n_estimators,<br/>max_samples,<br/>max_features,<br/>class_weight</p> | <ul style="list-style-type: none"> <li>- <i>n_estimators</i> spécifie le nombre d'estimateurs (par exemple, decision tree et logistque regression) utilisés dans l'ensemble</li> <li>- <i>max_samples</i> définit le nombre maximum d'échantillons utilisés pour entraîner chaque estimateur</li> <li>- <i>max_features</i> spécifie le nombre de caractéristiques considérées pour faire une sélection aléatoire à chaque nœud de décision</li> <li>- <i>class_weight</i> permet de gérer les inégalités de poids entre les différentes classes</li> </ul> |

TABLEAU A.2 - Description des algorithmes de ML utilisés et du rôle de certains de leurs paramètres clés [178–180] (Suite)

| Algorithme                        | Description   | Paramètres clés  | Rôles des paramètres   |
|-----------------------------------|---|--|--|
| RandomForest<br>(Forêt Aléatoire) | Random Forest est un algorithme d'ensemble. Il construit plusieurs arbres de décision à partir d'un échantillon aléatoire des données d'entraînement et un sous-ensemble aléatoire de caractéristiques. Il utilise la moyenne ou la majorité des prédictions pour améliorer la précision et la stabilité des prédictions. | n_estimators,<br>criterion,<br>max_depth,<br>min_samples_split,<br>max_features,<br>class_weight | - <i>n_estimators</i> définit le nombre d'arbres de décision utilisés dans l'ensemble<br>- <i>criterion</i> définit la mesure utilisée pour déterminer la qualité d'une division<br>- <i>max_depth</i> définit la profondeur maximale de l'arbre<br>- <i>min_samples_split</i> définit le nombre minimum d'échantillons requis pour effectuer une division<br>- <i>max_features</i> définit le nombre maximum de <i>features</i> utilisées pour faire une division |
| Adaboost                          | Adaboost est un algorithme d'ensemble. Il entraîne plusieurs modèles de faible complexité en attribuant un poids plus élevé aux observations mal classées pour aider les modèles suivants à mieux performer.  | n_estimators,<br>learning_rate,<br>class_weight  | - <i>n_estimators</i> définit le nombre de modèles utilisés dans l'ensemble<br>- <i>learning_rate</i> détermine la quantité de poids accordée à chaque estimateur<br>- <i>class_weight</i> permet de gérer les inégalités de poids entre les différentes classes   |

TABLEAU A.2 - Description des algorithmes de ML utilisés et du rôle de certains de leurs paramètres clés [178–180] (Suite)

| Algorithme       | Description   | Paramètres clés   | Rôles des paramètres   |
|------------------|---|---|--|
| GradientBoosting | Gradient Boosting est un algorithme d'ensemble. Il utilise une technique de régression pour améliorer les prédictions en ajoutant des modèles pour corriger les erreurs des modèles précédents. | n_estimators,<br>learning_rate,<br>max_depth,<br>min_samples_split,<br>min_samples_leaf,<br>subsample | - Les paramètres <i>n_estimators</i> , <i>learning_rate</i> , <i>max_depth</i> , <i>min_samples_split</i> , et <i>min_samples_leaf</i> fonctionnent de la même manière que<br>- <i>subsample</i> détermine la proportion d'échantillons utilisés pour entraîner chaque arbre |



TABLEAU A.2 - Description des algorithmes de ML utilisés et du rôle de certains de leurs paramètres clés [178–180] (Suite)

| Algorithme                             | Description  | Paramètres clés  | Rôles des paramètres   |
|--|--|--|--|
| Xgboost<br>(Extreme Gradient Boosting) | Xgboost est une implémentation de gradient boosting, optimisée pour la performance et la vitesse. Il entraîne plusieurs arbres de décision de manière itérative pour optimiser les poids des arbres et minimiser la perte. À chaque itération, les arbres précédents sont utilisés pour faire des prédictions, et les erreurs sont utilisées pour ajuster les poids des arbres suivants. | n_estimators,<br>learning_rate,<br>max_depth,<br>gamma,<br>reg_lambda,<br>reg_alpha,<br>scale_pos_weight | - <i>n_estimators</i> spécifie le nombre d'estimateurs (arbres) utilisés dans l'ensemble<br>- <i>learning_rate</i> détermine la quantité de poids accordée à chaque estimateur<br>- <i>max_depth</i> détermine la profondeur maximale des arbres<br>- <i>gamma</i> détermine la perte minimale requise pour effectuer une séparation sur un nœud<br>- <i>reg_lambda</i> est un paramètre de régularisation L2 pour les feuilles de l'arbre<br>- <i>reg_alpha</i> est un paramètre de régularisation L1 pour les feuilles de l'arbre<br>- <i>scale_pos_weight</i> utilisé pour pondérer les erreurs commises sur les observations négatives par rapport aux erreurs commises sur les observations positives |

TABLEAU A.2 - Description des algorithmes de ML utilisés et du rôle de certains de leurs paramètres clés [178–180] (Suite)

| Algorithme | Description  | Paramètres clés                    | Rôles des paramètres   |
|------------|--|------------------------------------|--|
| Stacking   | <p>Le Stacking est un algorithme d'ensemble. Il consiste à entraîner plusieurs modèles de base indépendamment et à les utiliser pour faire des prédictions sur un jeu de données de test. Les prédictions des modèles de base sont ensuite utilisées comme entrées pour un modèle de (métamodèle) qui fait la prédiction finale.</p> | <p>base models,<br/>meta-model</p> | <p>- <i>base models</i> spécifie la liste de modèles de base utilisés pour faire des prédictions intermédiaires<br/>- <i>meta-model</i> spécifie le modèle de métamodèle utilisé pour faire la prédiction finale</p> |

TABLEAU A.2 - Description des algorithmes de ML utilisés et du rôle de certains de leurs paramètres clés [178–180] (Suite)

| Algorithme | Description  | Paramètres clés                   | Rôles des paramètres  |
|------------|--|-----------------------------------|---|
| Voting     | <p>Voting est un algorithme d'ensemble. Il consiste à faire des prédictions en utilisant les prédictions de plusieurs modèles de machine learning et à les combiner en utilisant une stratégie de vote, comme la moyenne ou la majorité.</p> | <p>estimators, voting, weight</p> | <p>- <i>voting</i> spécifie la stratégie de vote (<i>hard</i> ou <i>soft</i>) à utiliser. Dans le <i>Hard voting</i> chaque modèle fait une prédiction indépendante et la classe finale prédite est celle avec la majorité des votes. Dans le <i>Soft voting</i> les probabilités prédites par chaque modèle sont combinées en utilisant une moyenne ou une pondération pour obtenir la prédiction finale. La pondération s'effectue au moyen du paramètre <i>weight</i>. Il peut donner une pondération spécifique à chaque modèle dans la combinaison de leurs probabilités.</p> <p>- <i>estimators</i> spécifie la liste de modèles de machine learning à inclure dans le vote</p> |

TABLEAU A.3 – Performance des modèles créés par la méthode *All Features* pour la prédiction de l'aggravation de l'EDSS

| Modèle                                | AUC score |       |       | F1 score |       |       |       |       |
|---------------------------------------|-----------|-------|-------|----------|-------|-------|-------|-------|
|                                       | max       | min   | mean  | std      | max   | min   | mean  | std   |
| LogisticRegression                    | 0.635     | 0.462 | 0.548 | 0.060    | 0.737 | 0.698 | 0.715 | 0.014 |
| DecisionTreeClassifier                | 0.548     | 0.418 | 0.473 | 0.043    | 0.777 | 0.582 | 0.677 | 0.069 |
| BaggingClassifier_LogisticRegression  | 0.628     | 0.460 | 0.532 | 0.061    | 0.792 | 0.751 | 0.767 | 0.014 |
| BaggingClassifier_DecisionTree        | 0.592     | 0.429 | 0.491 | 0.055    | 0.830 | 0.755 | 0.783 | 0.025 |
| AdaBoostClassifier_DecisionTree       | 0.513     | 0.429 | 0.472 | 0.030    | 0.810 | 0.755 | 0.780 | 0.018 |
| RandomForestClassifier                | 0.600     | 0.462 | 0.536 | 0.048    | 0.783 | 0.700 | 0.732 | 0.036 |
| GradientBoostingClassifier            | 0.500     | 0.486 | 0.496 | 0.005    | 0.905 | 0.883 | 0.899 | 0.008 |
| XGBoostClassifier                     | 0.637     | 0.427 | 0.523 | 0.067    | 0.815 | 0.769 | 0.788 | 0.016 |
| SVCClassifier                         | 0.672     | 0.434 | 0.559 | 0.081    | 0.728 | 0.631 | 0.671 | 0.031 |
| BaggingClassifier_SVC                 | 0.500     | 0.500 | 0.500 | 0.000    | 0.905 | 0.896 | 0.903 | 0.004 |
| AdaBoostClassifier_LogisticRegression | 0.615     | 0.464 | 0.555 | 0.054    | 0.731 | 0.636 | 0.709 | 0.037 |
| AdaBoostClassifier_SVC                | 0.500     | 0.500 | 0.500 | 0.000    | 0.905 | 0.896 | 0.903 | 0.004 |

TABLEAU A.3 - Performance des modèles créés par la méthode *All Features* pour la prédiction de l'aggravation de l'EDSS (Suite)

| Modèle                                | Vrai Positif |       |       |        | Vrai Négatif |     |      |       |
|---------------------------------------|--------------|-------|-------|--------|--------------|-----|------|-------|
|                                       | max          | min   | mean  | std    | max          | min | mean | std   |
| LogisticRegression                    | 96.0         | 89.0  | 91.6  | 2.417  | 7.0          | 3.0 | 4.8  | 1.327 |
| DecisionTreeClassifier                | 108.0        | 67.0  | 86.2  | 14.932 | 6.0          | 2.0 | 3.6  | 1.625 |
| BaggingClassifier_LogisticRegression  | 109.0        | 101.0 | 104.0 | 2.828  | 6.0          | 2.0 | 3.6  | 1.356 |
| BaggingClassifier_DecisionTree        | 119.0        | 102.0 | 109.2 | 5.636  | 4.0          | 1.0 | 2.4  | 1.020 |
| AdaBoostClassifier_DecisionTree       | 116.0        | 104.0 | 109.0 | 4.099  | 3.0          | 1.0 | 2.0  | 0.632 |
| RandomForestClassifier                | 107.0        | 87.0  | 96.0  | 8.295  | 6.0          | 3.0 | 4.2  | 0.980 |
| GradientBoostingClassifier            | 146.0        | 141.0 | 144.4 | 1.744  | 0.0          | 0.0 | 0.0  | 0.000 |
| XGBoostClassifier                     | 113.0        | 105.0 | 109.4 | 2.728  | 5.0          | 1.0 | 3.0  | 1.265 |
| SVCClassifier                         | 94.0         | 74.0  | 82.6  | 6.591  | 8.0          | 3.0 | 5.6  | 1.744 |
| BaggingCassitier_SVC                  | 146.0        | 145.0 | 145.6 | 0.490  | 0.0          | 0.0 | 0.0  | 0.000 |
| AdaBoostClassifier_LogisticRegression | 95.0         | 77.0  | 90.4  | 6.771  | 6.0          | 4.0 | 5.0  | 0.894 |
| AdaBoostClassifier_SVC                | 146.0        | 145.0 | 145.6 | 0.490  | 0.0          | 0.0 | 0.0  | 0.000 |

TABLEAU A.3 - Performance des modèles créés par la méthode *All Features* pour la prédiction de l'aggravation de l'EDSS (Suite)

| Modèle                                | Faux Positif |      |      |       | Faux Négatif |      |      |        |
|---------------------------------------|--------------|------|------|-------|--------------|------|------|--------|
|                                       | max          | min  | mean | std   | max          | min  | mean | std    |
| LogisticRegression                    | 7.0          | 4.0  | 5.4  | 1.020 | 57.0         | 50.0 | 54.0 | 2.366  |
| DecisionTreeClassifier                | 8.0          | 4.0  | 6.6  | 1.744 | 79.0         | 38.0 | 59.4 | 14.814 |
| BaggingClassifier_LogisticRegression  | 8.0          | 5.0  | 6.6  | 1.020 | 45.0         | 37.0 | 41.6 | 2.653  |
| BaggingClassifier_DecisionTree        | 9.0          | 7.0  | 7.8  | 0.748 | 44.0         | 26.0 | 36.4 | 6.020  |
| AdaBoostClassifier_DecisionTree       | 9.0          | 7.0  | 8.2  | 0.748 | 41.0         | 30.0 | 36.6 | 3.929  |
| RandomForestClassifier                | 7.0          | 4.0  | 6.0  | 1.095 | 58.0         | 39.0 | 49.6 | 8.309  |
| GradientBoostingClassifier            | 11.0         | 10.0 | 10.2 | 0.400 | 4.0          | 0.0  | 1.2  | 1.470  |
| XGBoostClassifier                     | 9.0          | 5.0  | 7.2  | 1.327 | 41.0         | 33.0 | 36.2 | 2.786  |
| SVCClassifier                         | 7.0          | 2.0  | 4.6  | 1.855 | 72.0         | 52.0 | 63.0 | 6.512  |
| BaggingClassifier_SVC                 | 11.0         | 10.0 | 10.2 | 0.400 | 0.0          | 0.0  | 0.0  | 0.000  |
| AdaBoostClassifier_LogisticRegression | 6.0          | 4.0  | 5.2  | 0.748 | 69.0         | 50.0 | 55.2 | 7.026  |
| AdaBoostClassifier_SVC                | 11.0         | 10.0 | 10.2 | 0.400 | 0.0          | 0.0  | 0.0  | 0.000  |

TABLEAU A.4 – Performance des modèles créés par la méthode *Physician Features* pour la prédiction de l'aggravation de l'EDSS

| Modèle                                | AUC score |       |       | F1 score |       |       |       |       |
|---------------------------------------|-----------|-------|-------|----------|-------|-------|-------|-------|
|                                       | max       | min   | mean  | std      | max   | min   | mean  | std   |
| LogisticRegression                    | 0.601     | 0.401 | 0.509 | 0.080    | 0.767 | 0.743 | 0.756 | 0.010 |
| Decision TreeClassifier               | 0.682     | 0.432 | 0.551 | 0.082    | 0.813 | 0.720 | 0.766 | 0.032 |
| BaggingClassifier_LogisticRegression  | 0.627     | 0.467 | 0.553 | 0.057    | 0.871 | 0.826 | 0.844 | 0.017 |
| BaggingClassifier_DecisionTree        | 0.685     | 0.432 | 0.552 | 0.084    | 0.817 | 0.720 | 0.767 | 0.034 |
| AdaBoostClassifier_DecisionTree       | 0.543     | 0.497 | 0.508 | 0.018    | 0.910 | 0.891 | 0.902 | 0.006 |
| RandomForestClassifier                | 0.572     | 0.474 | 0.521 | 0.035    | 0.834 | 0.789 | 0.817 | 0.017 |
| GradientBoostingClassifier            | 0.551     | 0.448 | 0.503 | 0.034    | 0.879 | 0.852 | 0.865 | 0.010 |
| XGBoostClassifier                     | 0.646     | 0.449 | 0.549 | 0.073    | 0.825 | 0.790 | 0.808 | 0.011 |
| SVCClassifier                         | 0.614     | 0.495 | 0.533 | 0.048    | 0.689 | 0.611 | 0.662 | 0.029 |
| BaggingClassifier_SVC                 | 0.500     | 0.500 | 0.500 | 0.000    | 0.904 | 0.894 | 0.902 | 0.004 |
| AdaBoostClassifier_LogisticRegression | 0.637     | 0.490 | 0.530 | 0.055    | 0.766 | 0.724 | 0.748 | 0.014 |
| AdaBoostClassifier SVC                | 0.500     | 0.500 | 0.500 | 0.000    | 0.904 | 0.894 | 0.902 | 0.004 |

TABLEAU A.4 - Performance des modèles créés par la méthode *Physician Features* pour la prédiction de l'aggravation de l'EDSS  
(Suite)

| Modèle                                | Vrai Positif |       |       |       | Vrai Négatif |     |      |       |
|---------------------------------------|--------------|-------|-------|-------|--------------|-----|------|-------|
|                                       | max          | min   | mean  | std   | max          | min | mean | std   |
| LogisticRegression                    | 104.0        | 98.0  | 101.0 | 1.897 | 5.0          | 1.0 | 3.2  | 1.600 |
| Decision TreeClassifier               | 110.0        | 92.0  | 102.4 | 7.310 | 6.0          | 1.0 | 4.0  | 1.789 |
| BaggingClassifier_LogisticRegression  | 128.0        | 119.0 | 122.2 | 3.187 | 4.0          | 1.0 | 2.6  | 1.020 |
| BaggingClassifier_DecisionTree        | 111.0        | 92.0  | 102.6 | 7.526 | 6.0          | 1.0 | 4.0  | 1.789 |
| AdaBoostClassifier_DecisionTree       | 144.0        | 142.0 | 143.2 | 0.980 | 1.0          | 0.0 | 0.2  | 0.400 |
| RandomForestClassifier                | 122.0        | 107.0 | 116.0 | 5.254 | 4.0          | 1.0 | 2.4  | 1.020 |
| GradientBoostingClassifier            | 134.0        | 128.0 | 130.6 | 2.154 | 2.0          | 0.0 | 1.0  | 0.632 |
| XGBoostClassifier                     | 115.0        | 106.0 | 112.8 | 3.429 | 5.0          | 1.0 | 3.2  | 1.600 |
| SVCClassifier                         | 86.0         | 71.0  | 80.2  | 5.776 | 7.0          | 4.0 | 5.2  | 1.166 |
| BaggingClassifier_SVC                 | 144.0        | 143.0 | 143.8 | 0.400 | 0.0          | 0.0 | 0.0  | 0.000 |
| AdaBoostClassifier_LogisticRegression | 103.0        | 93.0  | 98.6  | 3.611 | 6.0          | 3.0 | 3.8  | 1.166 |
| AdaBoostClassifier SVC                | 144.0        | 143.0 | 143.8 | 0.400 | 0.0          | 0.0 | 0.0  | 0.000 |



TABLEAU A.4 - Performance des modèles créés par la méthode *Physician Features* pour la prédiction de l'aggravation de l'EDSS  
(Suite)

| Modèle                                | Faux Positif |      |      | Faux Négatif |      |      |      |       |
|---------------------------------------|--------------|------|------|--------------|------|------|------|-------|
|                                       | max          | min  | mean | std          | max  | min  | mean | std   |
| LogisticRegression                    | 9.0          | 5.0  | 7.0  | 1.789        | 46.0 | 39.0 | 42.8 | 2.227 |
| Decision TreeClassifier               | 9.0          | 4.0  | 6.2  | 1.720        | 51.0 | 34.0 | 41.4 | 7.031 |
| BaggingClassifier_LogisticRegression  | 9.0          | 6.0  | 7.6  | 1.200        | 25.0 | 16.0 | 21.6 | 3.137 |
| BaggingClassifier_DecisionTree        | 9.0          | 4.0  | 6.2  | 1.720        | 51.0 | 33.0 | 41.2 | 7.250 |
| AdaBoostClassifier_DecisionTree       | 11.0         | 9.0  | 10.0 | 0.632        | 2.0  | 0.0  | 0.6  | 0.800 |
| RandomForestClassifier                | 9.0          | 6.0  | 7.8  | 0.980        | 37.0 | 22.0 | 27.8 | 5.307 |
| GradientBoostingClassifier            | 10.0         | 8.0  | 9.2  | 0.748        | 16.0 | 10.0 | 13.2 | 2.315 |
| XGBoostClassifier                     | 9.0          | 5.0  | 7.0  | 1.673        | 38.0 | 29.0 | 31.0 | 3.521 |
| SVCClassifier                         | 6.0          | 3.0  | 5.0  | 1.095        | 73.0 | 58.0 | 63.6 | 5.886 |
| BaggingClassifier_SVC                 | 11.0         | 10.0 | 10.2 | 0.400        | 0.0  | 0.0  | 0.0  | 0.000 |
| AdaBoostClassifier_LogisticRegression | 8.0          | 4.0  | 6.4  | 1.356        | 51.0 | 41.0 | 45.2 | 3.816 |
| AdaBoostClassifier_SVC                | 11.0         | 10.0 | 10.2 | 0.400        | 0.0  | 0.0  | 0.0  | 0.000 |

TABLEAU A.5 – Performance des modèles créés par la méthode *Physician Features & categories* pour la prédiction de l'aggravation de l'EDSS

| Modèle                                | AUC score |       |       | F1 score |       |       |       |       |
|---------------------------------------|-----------|-------|-------|----------|-------|-------|-------|-------|
|                                       | max       | min   | mean  | std      | max   | min   | mean  | std   |
| LogisticRegression                    | 0.615     | 0.491 | 0.561 | 0.049    | 0.711 | 0.626 | 0.679 | 0.028 |
| DecisionTreeClassifier                | 0.625     | 0.443 | 0.549 | 0.063    | 0.743 | 0.608 | 0.683 | 0.051 |
| BaggingClassifier_LogisticRegression  | 0.615     | 0.491 | 0.561 | 0.049    | 0.711 | 0.626 | 0.679 | 0.028 |
| BaggingClassifier_Decision            | 0.625     | 0.443 | 0.549 | 0.063    | 0.743 | 0.614 | 0.684 | 0.050 |
| AdaBoostClassifier Decision Tree      | 0.577     | 0.416 | 0.512 | 0.066    | 0.805 | 0.751 | 0.774 | 0.022 |
| RandomForestClassifier                | 0.660     | 0.429 | 0.523 | 0.080    | 0.811 | 0.742 | 0.776 | 0.024 |
| GradientBoostingClassifier            | 0.529     | 0.483 | 0.498 | 0.017    | 0.898 | 0.888 | 0.893 | 0.004 |
| XGBoostClassifier                     | 0.527     | 0.471 | 0.504 | 0.022    | 0.740 | 0.646 | 0.688 | 0.036 |
| SVCClassifier                         | 0.692     | 0.539 | 0.598 | 0.055    | 0.767 | 0.708 | 0.745 | 0.022 |
| BaggingClassifier_SVC                 | 0.500     | 0.500 | 0.500 | 0.000    | 0.905 | 0.896 | 0.903 | 0.004 |
| AdaBoostClassifier_LogisticRegression | 0.622     | 0.503 | 0.550 | 0.045    | 0.732 | 0.639 | 0.690 | 0.032 |
| AdaBoostClassifier_SVC                | 0.500     | 0.500 | 0.500 | 0.000    | 0.905 | 0.896 | 0.903 | 0.004 |

TABLEAU A.5 - Performance des modèles créés par la méthode *Physician Features & categories* pour la prédiction de l'aggravation de l'EDSS (Suite)

| Modèle                                | Vrai Positif |       |       | Vrai Négatif |     |     |      |       |
|---------------------------------------|--------------|-------|-------|--------------|-----|-----|------|-------|
|                                       | max          | min   | mean  | std          | max | min | mean | std   |
| LogisticRegression                    | 91.0         | 73.0  | 84.0  | 5.933        | 8.0 | 4.0 | 5.6  | 1.497 |
| DecisionTreeClassifier                | 98.0         | 70.0  | 85.4  | 10.288       | 7.0 | 3.0 | 5.2  | 1.327 |
| BaggingClassifier_LogisticRegression  | 91.0         | 73.0  | 84.0  | 5.933        | 8.0 | 4.0 | 5.6  | 1.497 |
| BaggingClassifier_Decision            | 98.0         | 71.0  | 85.6  | 9.992        | 7.0 | 3.0 | 5.2  | 1.327 |
| AdaBoostClassifier_Decision Tree      | 113.0        | 100.0 | 106.4 | 4.673        | 4.0 | 1.0 | 3.0  | 1.265 |
| RandomForestClassifier                | 116.0        | 98.0  | 106.6 | 6.053        | 6.0 | 1.0 | 3.2  | 1.720 |
| GradientBoostingClassifier            | 145.0        | 140.0 | 142.0 | 2.098        | 1.0 | 0.0 | 0.2  | 0.400 |
| XGBoostClassifier                     | 98.0         | 79.0  | 87.0  | 7.403        | 5.0 | 3.0 | 4.2  | 0.748 |
| SVCClassifier                         | 102.0        | 90.0  | 97.2  | 4.261        | 7.0 | 4.0 | 5.4  | 1.020 |
| BaggingClassifier_SVC                 | 146.0        | 145.0 | 145.6 | 0.490        | 0.0 | 0.0 | 0.0  | 0.000 |
| AdaBoostClassifier_LogisticRegression | 94.0         | 76.0  | 86.4  | 6.344        | 7.0 | 4.0 | 5.2  | 1.166 |
| AdaBoostClassifier_SVC                | 146.0        | 145.0 | 145.6 | 0.490        | 0.0 | 0.0 | 0.0  | 0.000 |

TABLEAU A.5 - Performance des modèles créés par la méthode *Physician Features & categories* pour la prédiction de l'aggravation de l'EDSS (Suite)

| Modèle                                | Faux Positif |      |      | Faux Négatif |      |      |      |        |
|---------------------------------------|--------------|------|------|--------------|------|------|------|--------|
|                                       | max          | min  | mean | std          | max  | min  | mean | std    |
| LogisticRegression                    | 6.0          | 3.0  | 4.6  | 1.200        | 72.0 | 54.0 | 61.6 | 5.817  |
| DecisionTreeClassifier                | 7.0          | 3.0  | 5.0  | 1.414        | 76.0 | 47.0 | 60.2 | 10.534 |
| BaggingClassifier_LogisticRegression  | 6.0          | 3.0  | 4.6  | 1.200        | 72.0 | 54.0 | 61.6 | 5.817  |
| BaggingClassifier_Decision            | 7.0          | 3.0  | 5.0  | 1.414        | 75.0 | 47.0 | 60.0 | 10.237 |
| AdaBoostClassifier_Decision Tree      | 9.0          | 6.0  | 7.2  | 1.166        | 46.0 | 32.0 | 39.2 | 4.956  |
| RandomForestClassifier                | 9.0          | 4.0  | 7.0  | 1.789        | 48.0 | 29.0 | 39.0 | 6.481  |
| GradientBoostingClassifier            | 11.0         | 9.0  | 10.0 | 0.632        | 6.0  | 0.0  | 3.6  | 2.245  |
| XGBoostClassifier                     | 7.0          | 5.0  | 6.0  | 0.632        | 67.0 | 47.0 | 58.6 | 7.710  |
| SVCClassifier                         | 6.0          | 3.0  | 4.8  | 0.980        | 56.0 | 43.0 | 48.4 | 4.409  |
| BaggingClassifier_SVC                 | 11.0         | 10.0 | 10.2 | 0.400        | 0.0  | 0.0  | 0.0  | 0.000  |
| AdaBoostClassifier_LogisticRegression | 6.0          | 4.0  | 5.0  | 0.894        | 69.0 | 52.0 | 59.2 | 6.079  |
| AdaBoostClassifier_SVC                | 11.0         | 10.0 | 10.2 | 0.400        | 0.0  | 0.0  | 0.0  | 0.000  |

TABLEAU A.6 – Performance des modèles créés par la méthode *All Features* pour la prédiction de l'aggravation des lésions

| Modèle                                | AUC score |       |       | F1 score |       |       |       |       |
|---------------------------------------|-----------|-------|-------|----------|-------|-------|-------|-------|
|                                       | max       | min   | mean  | std      | max   | min   | mean  | std   |
| LogisticRegression                    | 0.746     | 0.662 | 0.701 | 0.035    | 0.746 | 0.662 | 0.701 | 0.035 |
| DecisionTreeClassifier                | 0.752     | 0.591 | 0.687 | 0.058    | 0.751 | 0.591 | 0.686 | 0.058 |
| BaggingClassifier_LogisticRegression  | 0.746     | 0.669 | 0.704 | 0.033    | 0.746 | 0.668 | 0.704 | 0.033 |
| BaggingClassifier_Decision            | 0.753     | 0.642 | 0.698 | 0.044    | 0.753 | 0.641 | 0.698 | 0.044 |
| AdaBoostClassifier Decision Tree      | 0.773     | 0.648 | 0.714 | 0.050    | 0.773 | 0.647 | 0.713 | 0.051 |
| RandomForestClassifier                | 0.740     | 0.655 | 0.712 | 0.033    | 0.740 | 0.655 | 0.712 | 0.033 |
| GradientBoostingClassifier            | 0.760     | 0.656 | 0.709 | 0.042    | 0.760 | 0.656 | 0.709 | 0.042 |
| XGBoostClassifier                     | 0.779     | 0.656 | 0.723 | 0.051    | 0.779 | 0.656 | 0.723 | 0.051 |
| SVCClassifier                         | 0.753     | 0.649 | 0.713 | 0.034    | 0.753 | 0.649 | 0.713 | 0.034 |
| BaggingClassifier_SVC                 | 0.766     | 0.675 | 0.722 | 0.036    | 0.766 | 0.675 | 0.722 | 0.036 |
| AdaBoostClassifier_LogisticRegression | 0.766     | 0.642 | 0.708 | 0.047    | 0.766 | 0.642 | 0.707 | 0.047 |
| AdaBoostClassifier_SVC                | 0.760     | 0.643 | 0.714 | 0.051    | 0.760 | 0.643 | 0.714 | 0.051 |

TABLEAU A.6 - Performance des modèles créés par la méthode *All Features* pour la prédiction de l'aggravation des lésions (Suite)

| Modèle                                | Vrai Positif |      |      |       | Vrai Négatif |      |      |       |
|---------------------------------------|--------------|------|------|-------|--------------|------|------|-------|
|                                       | max          | min  | mean | std   | max          | min  | mean | std   |
| LogisticRegression                    | 61.0         | 48.0 | 54.2 | 4.707 | 55.0         | 51.0 | 53.8 | 1.470 |
| DecisionTreeClassifier                | 65.0         | 47.0 | 54.0 | 6.573 | 61.0         | 44.0 | 51.8 | 5.636 |
| BaggingClassifier_LogisticRegression  | 61.0         | 49.0 | 54.0 | 4.490 | 55.0         | 51.0 | 53.0 | 1.356 |
| BaggingClassifier_Decision            | 62.0         | 56.0 | 58.4 | 2.245 | 56.0         | 43.0 | 49.2 | 4.665 |
| AdaBoostClassifier Decision Tree      | 60.0         | 56.0 | 58.4 | 1.625 | 59.0         | 43.0 | 51.6 | 6.437 |
| RandomForestClassifier                | 59.0         | 54.0 | 56.6 | 1.855 | 57.0         | 47.0 | 53.0 | 3.406 |
| GradientBoostingClassifier            | 60.0         | 50.0 | 56.2 | 3.544 | 57.0         | 47.0 | 53.0 | 3.795 |
| XGBoostClassifier                     | 61.0         | 51.0 | 57.0 | 4.147 | 59.0         | 48.0 | 54.4 | 3.929 |
| SVCClassifier                         | 60.0         | 51.0 | 55.8 | 3.059 | 56.0         | 49.0 | 54.0 | 2.608 |
| BaggingClassifier_SVC                 | 62.0         | 52.0 | 57.4 | 4.128 | 56.0         | 52.0 | 53.8 | 1.600 |
| AdaBoostClassifier_LogisticRegression | 61.0         | 50.0 | 56.6 | 4.224 | 57.0         | 45.0 | 52.4 | 4.454 |
| AdaBoostClassifier_SVC                | 61.0         | 48.0 | 55.0 | 5.404 | 58.0         | 51.0 | 55.0 | 2.757 |

TABLEAU A.6 - Performance des modèles créés par la méthode *All Features* pour la prédiction de l'aggravation des lésions (Suite)

| Modèle                                | Faux Positif |      |      |       | Faux Négatif |      |      |        |
|---------------------------------------|--------------|------|------|-------|--------------|------|------|--------|
|                                       | max          | min  | mean | std   | max          | min  | mean | std    |
| LogisticRegression                    | 25.0         | 21.0 | 22.2 | 1.470 | 30.0         | 17.0 | 23.8 | 4.707  |
| DecisionTreeClassifier                | 32.0         | 15.0 | 24.2 | 5.636 | 31.0         | 13.0 | 24.0 | 6.573  |
| BaggingClassifier_LogisticRegression  | 25.0         | 21.0 | 22.4 | 1.356 | 29.0         | 17.0 | 23.2 | 4.4490 |
| BaggingClassifier_Decision            | 33.0         | 20.0 | 26.8 | 4.665 | 22.0         | 16.0 | 19.6 | 2.245  |
| AdaBoostClassifier Decision Tree      | 33.0         | 17.0 | 24.4 | 6.437 | 22.0         | 18.0 | 19.6 | 1.625  |
| RandomForestClassifier                | 29.0         | 19.0 | 23.0 | 3.406 | 24.0         | 19.0 | 21.4 | 1.855  |
| GradientBoostingClassifier            | 29.0         | 19.0 | 23.0 | 3.795 | 28.0         | 18.0 | 21.8 | 3.544  |
| XGBoostClassifier                     | 28.0         | 17.0 | 21.6 | 3.929 | 27.0         | 17.0 | 21.0 | 4.147  |
| SVCClassifier                         | 27.0         | 20.0 | 22.0 | 2.608 | 27.0         | 18.0 | 22.2 | 3.059  |
| BaggingClassifier_SVC                 | 24.0         | 20.0 | 22.2 | 1.600 | 26.0         | 16.0 | 20.6 | 4.128  |
| AdaBoostClassifier_LogisticRegression | 31.0         | 19.0 | 23.6 | 4.454 | 28.0         | 17.0 | 21.4 | 4.224  |
| AdaBoostClassifier_SVC                | 25.0         | 18.0 | 21.0 | 2.757 | 30.0         | 17.0 | 23.0 | 5.404  |

TABLEAU A.7 – Performance des modèles créés par la méthode *Physician Features* pour la prédiction de l'aggravation des lésions

| Modèle                                | AUC score |       |       | F1 score |       |       |       |       |
|---------------------------------------|-----------|-------|-------|----------|-------|-------|-------|-------|
|                                       | max       | min   | mean  | std      | max   | min   | mean  | std   |
| LogisticRegression                    | 0.761     | 0.699 | 0.729 | 0.026    | 0.761 | 0.699 | 0.729 | 0.026 |
| DecisionTreeClassifier                | 0.738     | 0.635 | 0.701 | 0.038    | 0.737 | 0.634 | 0.700 | 0.039 |
| BaggingClassifier_LogisticRegression  | 0.763     | 0.686 | 0.724 | 0.030    | 0.763 | 0.686 | 0.724 | 0.030 |
| BaggingClassifier_Decision            | 0.768     | 0.659 | 0.708 | 0.040    | 0.768 | 0.657 | 0.708 | 0.041 |
| AdaBoostClassifier Decision Tree      | 0.750     | 0.666 | 0.706 | 0.032    | 0.750 | 0.666 | 0.705 | 0.032 |
| RandomForestClassifier                | 0.743     | 0.659 | 0.707 | 0.039    | 0.743 | 0.658 | 0.706 | 0.039 |
| GradientBoostingClassifier            | 0.749     | 0.659 | 0.708 | 0.033    | 0.749 | 0.658 | 0.708 | 0.034 |
| XGBoostClassifier                     | 0.750     | 0.673 | 0.715 | 0.030    | 0.750 | 0.673 | 0.714 | 0.030 |
| SVCClassifier                         | 0.763     | 0.673 | 0.724 | 0.036    | 0.763 | 0.673 | 0.724 | 0.036 |
| BaggingClassifier_SVC                 | 0.757     | 0.685 | 0.724 | 0.029    | 0.756 | 0.685 | 0.724 | 0.029 |
| AdaBoostClassifier_LogisticRegression | 0.768     | 0.673 | 0.720 | 0.035    | 0.768 | 0.673 | 0.720 | 0.035 |
| AdaBoostClassifier_SVC                | 0.761     | 0.685 | 0.721 | 0.030    | 0.761 | 0.685 | 0.721 | 0.030 |



TABLEAU A.7 - Performance des modèles créés par la méthode *Physician Features* pour la prédiction de l'aggravation des lésions  
(Suite)

| Modèle                                | Vrai Positif |      |      |       | Vrai Négatif |      |      |       |
|---------------------------------------|--------------|------|------|-------|--------------|------|------|-------|
|                                       | max          | min  | mean | std   | max          | min  | mean | std   |
| LogisticRegression                    | 61.0         | 54.0 | 57.8 | 2.561 | 60.0         | 52.0 | 55.8 | 2.993 |
| DecisionTreeClassifier                | 63.0         | 46.0 | 54.6 | 7.003 | 60.0         | 51.0 | 54.6 | 4.030 |
| BaggingClassifier_LogisticRegression  | 61.0         | 53.0 | 57.6 | 2.800 | 60.0         | 51.0 | 55.2 | 3.655 |
| BaggingClassifier_Decision            | 61.0         | 55.0 | 58.8 | 2.135 | 59.0         | 43.0 | 51.6 | 5.783 |
| AdaBoostClassifier_Decision Tree      | 64.0         | 54.0 | 58.4 | 3.611 | 56.0         | 48.0 | 51.6 | 2.728 |
| RandomForestClassifier                | 65.0         | 57.0 | 59.8 | 2.993 | 58.0         | 44.0 | 50.4 | 5.314 |
| GradientBoostingClassifier            | 65.0         | 55.0 | 58.8 | 3.370 | 55.0         | 45.0 | 51.6 | 3.666 |
| XGBoostClassifier                     | 64.0         | 53.0 | 58.6 | 3.980 | 58.0         | 46.0 | 52.8 | 4.534 |
| SVCClassifier                         | 62.0         | 52.0 | 57.0 | 3.406 | 61.0         | 50.0 | 54.8 | 4.308 |
| BaggingClassifier_SVC                 | 62.0         | 54.0 | 58.0 | 2.608 | 60.0         | 50.0 | 54.8 | 3.868 |
| AdaBoostClassifier_LogisticRegression | 62.0         | 54.0 | 57.2 | 3.059 | 60.0         | 50.0 | 55.0 | 4.290 |
| AdaBoostClassifier_SVC                | 63.0         | 55.0 | 59.0 | 3.033 | 59.0         | 49.0 | 53.4 | 3.720 |

TABLEAU A.7 - Performance des modèles créés par la méthode *Physician Features* pour la prédiction de l'aggravation des lésions  
(Suite)

| Modèle                                | Faux Positif |      |      | Faux Négatif |      |      |      |       |
|---------------------------------------|--------------|------|------|--------------|------|------|------|-------|
|                                       | max          | min  | mean | std          | max  | min  | mean | std   |
| LogisticRegression                    | 25.0         | 17.0 | 21.2 | 2.993        | 25.0 | 18.0 | 21.0 | 2.757 |
| DecisionTreeClassifier                | 26.0         | 17.0 | 22.4 | 4.030        | 33.0 | 16.0 | 24.2 | 7.194 |
| BaggingClassifier_LogisticRegression  | 26.0         | 17.0 | 21.8 | 3.655        | 26.0 | 18.0 | 21.2 | 2.926 |
| BaggingClassifier_Decision            | 34.0         | 18.0 | 25.4 | 5.783        | 24.0 | 18.0 | 20.0 | 2.280 |
| AdaBoostClassifier_Decision Tree      | 29.0         | 21.0 | 25.4 | 2.728        | 25.0 | 15.0 | 20.4 | 3.555 |
| RandomForestClassifier                | 33.0         | 19.0 | 26.6 | 5.314        | 22.0 | 14.0 | 19.0 | 2.828 |
| GradientBoostingClassifier            | 32.0         | 22.0 | 25.4 | 3.666        | 24.0 | 14.0 | 20.0 | 3.286 |
| XGBoostClassifier                     | 31.0         | 19.0 | 24.2 | 4.534        | 26.0 | 15.0 | 20.2 | 3.970 |
| SVCClassifier                         | 27.0         | 16.0 | 21.2 | 4.308        | 27.0 | 17.0 | 21.8 | 3.429 |
| BaggingClassifier_SVC                 | 27.0         | 17.0 | 22.2 | 3.868        | 25.0 | 17.0 | 20.8 | 2.713 |
| AdaBoostClassifier_LogisticRegression | 27.0         | 17.0 | 22.0 | 4.290        | 25.0 | 17.0 | 21.6 | 3.200 |
| AdaBoostClassifier_SVC                | 28.0         | 18.0 | 23.6 | 3.720        | 24.0 | 16.0 | 19.8 | 3.250 |

TABLEAU A.8 – Performance des modèles créés par la méthode *Physician Features & categories* pour la prédiction de l'aggravation des lésions

| Modèle                                | AUC score |       |       | F1 score |       |       |       |       |
|---------------------------------------|-----------|-------|-------|----------|-------|-------|-------|-------|
|                                       | max       | min   | mean  | std      | max   | min   | mean  | std   |
| LogisticRegression                    | 0.730     | 0.634 | 0.692 | 0.036    | 0.730 | 0.633 | 0.691 | 0.036 |
| DecisionTreeClassifier                | 0.749     | 0.629 | 0.699 | 0.047    | 0.749 | 0.627 | 0.698 | 0.047 |
| BaggingClassifier_LogisticRegression  | 0.730     | 0.634 | 0.695 | 0.037    | 0.730 | 0.632 | 0.695 | 0.037 |
| BaggingClassifier_Decision            | 0.749     | 0.640 | 0.707 | 0.045    | 0.749 | 0.640 | 0.707 | 0.045 |
| AdaBoostClassifier_Decision Tree      | 0.756     | 0.660 | 0.699 | 0.037    | 0.756 | 0.660 | 0.699 | 0.037 |
| RandomForestClassifier                | 0.774     | 0.679 | 0.725 | 0.035    | 0.774 | 0.679 | 0.725 | 0.035 |
| GradientBoostingClassifier            | 0.743     | 0.672 | 0.716 | 0.030    | 0.743 | 0.672 | 0.716 | 0.031 |
| XGBoostClassifier                     | 0.742     | 0.659 | 0.709 | 0.032    | 0.742 | 0.658 | 0.709 | 0.032 |
| SVCClassifier                         | 0.743     | 0.660 | 0.702 | 0.029    | 0.744 | 0.660 | 0.702 | 0.029 |
| BaggingClassifier_SVC                 | 0.744     | 0.685 | 0.707 | 0.022    | 0.744 | 0.685 | 0.707 | 0.022 |
| AdaBoostClassifier_LogisticRegression | 0.748     | 0.608 | 0.684 | 0.048    | 0.748 | 0.608 | 0.684 | 0.048 |
| AdaBoostClassifier_SVC                | 0.742     | 0.603 | 0.673 | 0.045    | 0.742 | 0.603 | 0.673 | 0.045 |

TABLEAU A.8 - Performance des modèles créés par la méthode *Physician Features & categories* pour la prédiction de l'aggravation des lésions (Suite)

| Modèle                                | Vrai Positif |      |      | Vrai Négatif |      |      |      |       |
|---------------------------------------|--------------|------|------|--------------|------|------|------|-------|
|                                       | max          | min  | mean | std          | max  | min  | mean | std   |
| LogisticRegression                    | 62.0         | 54.0 | 57.4 | 3.382        | 54.0 | 44.0 | 50.4 | 3.441 |
| DecisionTreeClassifier                | 64.0         | 46.0 | 57.2 | 7.985        | 54.0 | 49.0 | 51.8 | 1.720 |
| BaggingClassifier_LogisticRegression  | 62.0         | 55.0 | 58.0 | 2.608        | 54.0 | 43.0 | 50.4 | 3.929 |
| BaggingClassifier_Decision            | 63.0         | 53.0 | 58.6 | 3.878        | 55.0 | 45.0 | 51.6 | 3.555 |
| AdaBoostClassifier_Decision Tree      | 64.0         | 50.0 | 57.2 | 5.115        | 55.0 | 48.0 | 51.8 | 2.561 |
| RandomForestClassifier                | 62.0         | 56.0 | 59.6 | 2.939        | 58.0 | 50.0 | 53.4 | 2.653 |
| GradientBoostingClassifier            | 63.0         | 56.0 | 59.8 | 2.482        | 54.0 | 47.0 | 51.8 | 2.638 |
| XGBoostClassifier                     | 62.0         | 55.0 | 59.2 | 2.482        | 55.0 | 45.0 | 51.4 | 3.499 |
| SVCClassifier                         | 60.0         | 54.0 | 57.6 | 2.332        | 56.0 | 49.0 | 51.8 | 2.400 |
| BaggingClassifier_SVC                 | 60.0         | 57.0 | 58.4 | 1.020        | 57.0 | 49.0 | 51.8 | 2.786 |
| AdaBoostClassifier_LogisticRegression | 61.0         | 52.0 | 56.0 | 4.147        | 55.0 | 43.0 | 50.6 | 4.409 |
| AdaBoostClassifier_SVC                | 59.0         | 47.0 | 53.0 | 4.000        | 56.0 | 47.0 | 51.8 | 3.600 |

TABLEAU A.8 - Performance des modèles créés par la méthode *Physician Features & categories* pour la prédiction de l'aggravation des lésions (Suite)

| Modèle                                | Faux Positif |      |      | Faux Négatif |      |      |      |       |
|---------------------------------------|--------------|------|------|--------------|------|------|------|-------|
|                                       | max          | min  | mean | std          | max  | min  | mean | std   |
| LogisticRegression                    | 33.0         | 23.0 | 26.6 | 3.441        | 25.0 | 17.0 | 21.4 | 3.611 |
| DecisionTreeClassifier                | 28.0         | 23.0 | 25.2 | 1.720        | 33.0 | 15.0 | 21.6 | 8.139 |
| BaggingClassifier_LogisticRegression  | 34.0         | 23.0 | 26.6 | 3.929        | 24.0 | 17.0 | 20.8 | 2.786 |
| BaggingClassifier_Decision            | 32.0         | 22.0 | 25.4 | 3.555        | 26.0 | 16.0 | 20.2 | 4.020 |
| AdaBoostClassifier_Decision Tree      | 29.0         | 22.0 | 25.2 | 2.561        | 29.0 | 15.0 | 21.6 | 5.276 |
| RandomForestClassifier                | 27.0         | 19.0 | 23.6 | 2.653        | 23.0 | 16.0 | 19.2 | 3.124 |
| GradientBoostingClassifier            | 30.0         | 23.0 | 25.2 | 2.638        | 23.0 | 16.0 | 19.0 | 2.608 |
| XGBoostClassifier                     | 32.0         | 22.0 | 25.6 | 3.499        | 24.0 | 17.0 | 19.6 | 2.577 |
| SVCClassifier                         | 28.0         | 21.0 | 25.2 | 2.400        | 25.0 | 18.0 | 21.2 | 2.561 |
| BaggingClassifier_SVC                 | 28.0         | 20.0 | 25.2 | 2.786        | 22.0 | 18.0 | 20.4 | 1.356 |
| AdaBoostClassifier_LogisticRegression | 34.0         | 22.0 | 26.4 | 4.409        | 27.0 | 17.0 | 22.8 | 4.400 |
| AdaBoostClassifier_SVC                | 30.0         | 21.0 | 25.2 | 3.600        | 32.0 | 19.0 | 25.8 | 4.308 |



# **ANNEXE B : DOCUMENTS UTILISÉS DANS LES ÉVALUATIONS DE MS VISTA**

## Résultats du questionnaire d'évaluation de MS Vista pour le développement centré sur l'utilisateur

1. Le système optimise le bien-être de l'utilisateur (ergonomie, facilité d'usage, etc.).

|                               |   |
|-------------------------------|---|
| ● Tout à fait d'accord        | 0 |
| ● D'accord                    | 3 |
| ● Ni d'accord ni pas d'accord | 1 |
| ● Pas d'accord                | 0 |
| ● Pas du tout d'accord        | 0 |



2. Les calculs et les méthodes d'alertes sont automatiques et à la demande.

|                               |   |
|-------------------------------|---|
| ● Tout à fait d'accord        | 0 |
| ● D'accord                    | 3 |
| ● Ni d'accord ni pas d'accord | 1 |
| ● Pas d'accord                | 0 |
| ● Pas du tout d'accord        | 0 |



3. Les paramètres présentés dans le système sont correctement liés avec la médecine factuelle.

|                               |   |
|-------------------------------|---|
| ● Tout à fait d'accord        | 0 |
| ● D'accord                    | 1 |
| ● Ni d'accord ni pas d'accord | 3 |
| ● Pas d'accord                | 0 |
| ● Pas du tout d'accord        | 0 |





4. Les critères de sélection de la population de référence (POR) sont-ils suffisants pour contextualiser un patient ?

|                          |   |
|--------------------------|---|
| ● Largement suffisants   | 1 |
| ● Suffisants             | 0 |
| ● Ne sait pas            | 1 |
| ● Insuffisants           | 2 |
| ● Pas du tout suffisants | 0 |



5. Quelle est la probabilité que vous utilisiez cet outil pour prendre des décisions cliniques ?

|                        |   |
|------------------------|---|
| ● Fortement probable   | 0 |
| ● Modérément probable  | 1 |
| ● Ne sait pas          | 3 |
| ● Peu probable         | 0 |
| ● Pas du tout probable | 0 |



6. Dans l'ensemble, quel est votre niveau de satisfaction concernant notre application ?

|                       |   |
|-----------------------|---|
| ● Très satisfait(e)   | 0 |
| ● Satisfait(e)        | 4 |
| ● Peu satisfait(e)    | 0 |
| ● Peu insatisfait(e)  | 0 |
| ● Très insatisfait(e) | 0 |



7. À qui d'autres pensez-vous que cet outil pourrait être utile ?

- L'outil sera utile aux neurologues dans la relation avec le patient
- Principalement aux neurologues
- L'outil sera surtout pour le neurologue non expert et le patient

8. Comment pensez-vous que cette application peut s'améliorer ?

- Il faut absolument tenir compte des dates des événements et de la temporalité de cette maladie de plus, il faudrait annoncer quelles sont les données utilisées en réf et donner quelques caractéristiques de présentation dans l'idéal, il faudrait ajouter d'autres paramètres, non cliniques ou IRM comme tabac, génétique, obésité etc qui

sont aussi des facteurs pronostiques => ils sont rarement disponibles et c'est une limite qu'il faut assumer

- Données d'imageries à un temps donné +++ cf remarques ce jour
- En travaillant ensemble +++++
- Y travailler encore ensemble

9. Si nous appliquons vos suggestions pour améliorer l'application. Quelle est la probabilité que vous utilisiez cet outil pour prendre des décisions cliniques ?

|                        |   |
|------------------------|---|
| ● Fortement probable   | 0 |
| ● Modérément probable  | 1 |
| ● Ne sait pas          | 3 |
| ● Peu probable         | 0 |
| ● Pas du tout probable | 0 |



# Evaluation de MS Vista comme outil support de la communication médecin-patient

## Informations sur les participants – formulaire médecin

1. **Etes-vous :**

Femme     Homme

2. **Quelle est votre spécialité ?**

Neurologue     Autre (à préciser) .....

3. **Depuis combien d'années pratiquez-vous cette spécialité ?**

Moins de 5 ans (internat compris)

Entre 5 et 10 ans

Entre 11 et 20 ans

Plus que 20 ans

4. **Combien de patients atteints de sclérose en plaques recevez-vous par mois ?**

Moins de 6 patients

Entre 6 et 12 patients

Entre 13 et 25 patients

Plus que 25 patients

5. **Quel est votre attirance pour les outils informatiques ?**

1    2    3    4    5    6    7

Faible

Forte

6. **Avez-vous déjà utilisé un (des) système (s) d'aide à la décision clinique dans la sclérose en plaques ?**

Oui

Non

***Si oui, quel (s) outil (s) utilisez-vous ?***

.....  
.....

***A quelle fréquence utilisez-vous cet (ces) outil (s) ?***

Très fréquemment (*Exemple ; pour chaque patient lors de toutes les consultations*)

Fréquemment (*Exemple ; pour la plus part des patients en consultation*)

Parfois (*Exemple ; quelques fois par jour*)

Rarement (*Exemple ; moins souvent*)

Très rarement (*Exemple ; j'ai testé quelques fois*)





# Evaluation de MS Vista comme outil support de la communication médecin-patient

Mesure de l'utilisabilité du système – formulaire médecin (2)

Nous vous remerciant de votre participation 😊

**Gardez à l'esprit qu'il n'y a pas de bonne ou de mauvaise réponse. Seule votre opinion compte !**

## **1. MS Vista a amélioré globalement vos interactions avec le patient pendant cette consultation.**

- Tout à fait d'accord
- D'accord
- Ni d'accord ni pas d'accord
- Pas d'accord
- Pas du tout d'accord

## **2. MS Vista a amélioré la compréhension du patient.**

- Tout à fait d'accord
- D'accord
- Ni d'accord ni pas d'accord
- Pas d'accord
- Pas du tout d'accord

## **3. Le patient était satisfait de vos échanges autour de MS Vista.**

- Tout à fait d'accord
- D'accord
- Ni d'accord ni pas d'accord
- Pas d'accord
- Pas du tout d'accord

## **4. MS Vista peut améliorer l'adhésion du patient au traitement.**

- Tout à fait d'accord
- D'accord
- Ni d'accord ni pas d'accord
- Pas d'accord
- Pas du tout d'accord

**5. Les informations montrées par MS Vista sont toutes bonnes à montrer au patient.**

- Tout à fait d'accord
- D'accord
- Ni d'accord ni pas d'accord
- Pas d'accord
- Pas du tout d'accord

**6. Pensez-vous que MS Vista présente ; Trop peu, Assez ou Trop d'informations au patient ?**

**7. Comment pensez-vous que cette application peut être améliorée ? Merci de votre opinion.**

# Evaluation de MS Vista comme outil support de la communication médecin-patient

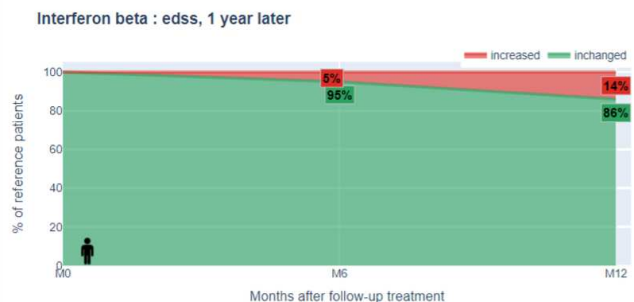
## Mesure de l'utilisabilité du système – formulaire patient

Nous vous remerciant de votre participation 😊

**Gardez à l'esprit qu'il n'y a pas de bonne ou de mauvaise réponse. Seule votre opinion compte !**

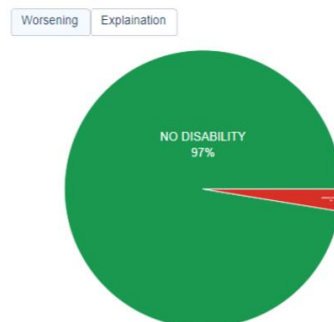
**1. Le graphique ci-dessous vous a aidé à comprendre l'évolution possible de votre score EDSS dans l'année qui va suivre votre visite.**

- Tout à fait d'accord
- D'accord
- Ni d'accord ni pas d'accord
- Pas d'accord
- Pas du tout d'accord



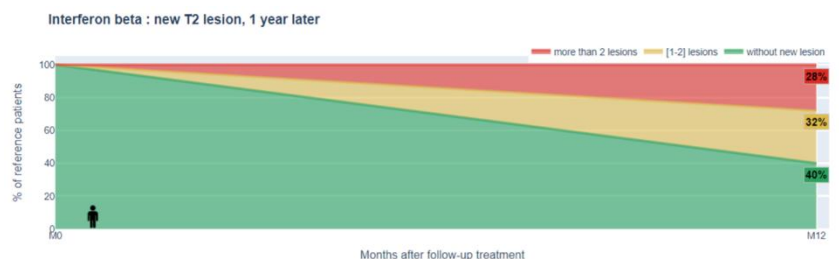
**2. Le graphique ci-dessous vous a aidé à comprendre la probabilité de développer un handicap dans l'année qui va suivre votre visite.**

- Tout à fait d'accord
- D'accord
- Ni d'accord ni pas d'accord
- Pas d'accord
- Pas du tout d'accord



**3. Le graphique ci-dessous vous a aidé à comprendre les probabilités de développer de nouvelles lésions sur votre IRM dans l'année qui va suivre votre visite.**

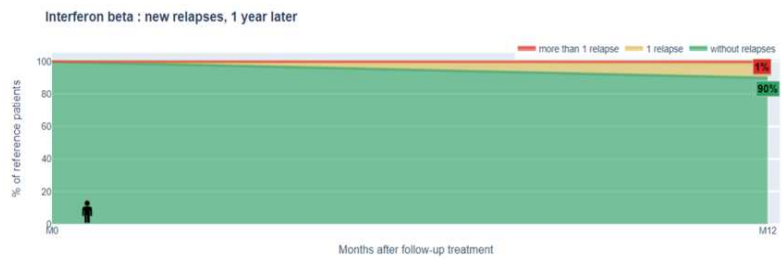
- Tout à fait d'accord
- D'accord
- Ni d'accord ni pas d'accord
- Pas d'accord
- Pas du tout d'accord





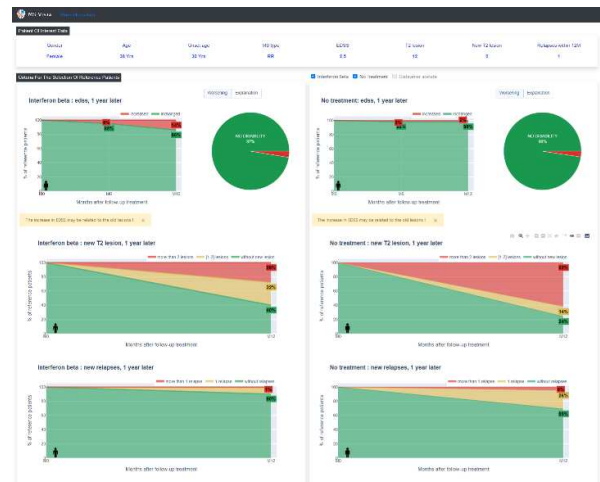
**4. Le graphique ci-dessous vous a aidé à comprendre les probabilités de développer de nouvelles poussées dans l'année qui va suivre votre visite.**

- Tout à fait d'accord
- D'accord
- Ni d'accord ni pas d'accord
- Pas d'accord
- Pas du tout d'accord



**5. Le graphique ci-dessous vous a convaincu d'adhérer au traitement proposé par votre médecin en observant son effet sur d'autres patients présentant des caractéristiques cliniques et démographiques similaires aux vôtres.**

- Tout à fait d'accord
- D'accord
- Ni d'accord ni pas d'accord
- Pas d'accord
- Pas du tout d'accord



**6. Vous pensez que l'utilisation de cette application vous a aidé dans votre communication avec votre médecin.**

- Tout à fait d'accord
- D'accord
- Ni d'accord ni pas d'accord
- Pas d'accord
- Pas du tout d'accord

**7. Comment pensez-vous que cette application peut s'améliorer ? Merci de votre opinion.**





**Titre :** Dialogue humain machine pour l'aide à la décision médicale

**Mots clés :** Interaction Humain-Machine, Intelligence Hybride, Apprentissage Automatique, Système d'Aide à la Décision Clinique, Collaboration Humain-Algorithmme, Sclérose En Plaques.

**Résumé :** Grâce aux progrès scientifiques et technologiques, la Médecine de Précision (MP) est désormais prometteuse pour la personnalisation de la décision médicale. L'objectif est de fournir des mesures spécifiques pour une prise en charge individualisée du patient. Les solutions développées à cette fin sont nombreuses, cependant, leur adoption dans les services de santé est limitée par des enjeux liés en particulier aux données et aux méthodes sous-jacentes.

Cette thèse s'intéresse aux mécanismes du dialogue humain-machine à travers la visualisation et l'interaction avec les données et les algorithmes pour fournir des solutions de MP raisonnables et compréhensibles par l'humain. En collaboration avec des neurologues, nous avons conçu et développé un système d'aide à la décision clinique appelé MS-Vista, fondé sur les principes de la MP.

MS-Vista offre une interface humain-machine d'interaction avec les données permettant de contextualiser les patients atteints de sclérose en plaques. Il permet d'examiner la qualité et la quantité des données. Il fournit également des visuels de projection de données illustrant les avantages potentiels de chaque option thérapeutique. Il s'agit d'aider à la personnalisation de la décision et à la communication médecin-patient. Les retours positifs fournis par les neurologues, notamment du fait que MS-Vista intègre leur raisonnement, nous a conduit par la suite à concevoir une approche hybride de prédiction basée sur la collaboration humain-algorithme. En impliquant les médecins dans le processus d'apprentissage, cette approche permet d'améliorer la performance du modèle de prédiction en combinant l'intelligence humaine et artificielle.

**Title :** Human computer dialogue for medical decision support

**Keywords :** Human-Computer Interaction, Hybrid Intelligence, Machine Learning, Clinical Decision Support System, Human-Algorithm Collaboration, Multiple Sclerosis.

**Abstract :** Given scientific and technological advances, Precision Medicine (PM) now holds promise for personalizing medical decision making. The aim is to provide specific measures for individualized patient management. Many solutions have been developed for this purpose, however, their adoption in healthcare services is limited by issues related especially to the underlying data and methods.

This thesis focuses on the mechanisms of human-computer dialog through visualization and interaction with data and algorithms to provide sensible and human-understandable PM solutions. In collaboration with neurologists, we designed a clinical decision support system called MS-Vista, based on PM principles.

MS-Vista provides a human-computer interface for data interaction to contextualize patients with multiple sclerosis. It enables the evaluation of data quality and quantity. It also provides interactive data projection illustrating the potential benefits of each therapeutic option. It can assist either in supporting the decision, either for the physician-patient communication. The positive feedback provided by neurologists, especially as MS-Vista integrates their reasoning, led us to go further and design a new hybrid prediction approach based on human-algorithm collaboration. By involving physicians in the learning process, this approach improves the performance of the prediction model through the combination of human and artificial intelligences.