



HAL
open science

Titre : Intelligence artificielle et lombalgie chronique : une approche multidimensionnelle pour personnaliser les soins de rééducation.

Steven Crombecque, Isabelle Caby, Claire Toulotte

► **To cite this version:**

Steven Crombecque, Isabelle Caby, Claire Toulotte. Titre : Intelligence artificielle et lombalgie chronique : une approche multidimensionnelle pour personnaliser les soins de rééducation.. 2025. <hal-04936750>

HAL Id: hal-04936750

<https://hal.science/hal-04936750v1>

Preprint submitted on 9 Feb 2025

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.



HAL Authorization

Titre : Intelligence artificielle et lombalgie chronique : une approche multidimensionnelle pour personnaliser les soins de rééducation.

Crombecque, Steven ^{a,b,c*} ; Caby, Isabelle ^a (ORCID : [0000-0002-9607-2177](https://orcid.org/0000-0002-9607-2177)) ; Toulotte, Claire ^a (ORCID : [0000-0002-5985-396X](https://orcid.org/0000-0002-5985-396X))

Affiliations des auteurs

- a. Univ. Artois, Univ. Lille, Univ. Littoral Côte d'Opale, ULR 7369 - URePSSS - Unité de Recherche Pluridisciplinaire Sport Santé Société, F-62800 Liévin, France
- b. Fondation Hopale, 62608 Berck, France.
- c. Centre Hélène-Borel, 59283 Raimbeaucourt, France

* Correspondance :

- Crombecque Steven
- 17 rue Louis Dupont, 59450 Sin Le Noble, France
- 06 20 29 54 10
- Stevencrombecque@univ-artois.fr
- Numéro ORCID : [0009-0003-5429-5887](https://orcid.org/0009-0003-5429-5887)

Nombre de pages : 31

Nombre de figures et tableaux : 1 figure et 2 tableaux

Résumé

Introduction

La lombalgie, touchant 83 % de la population française au cours de leur vie, représente un

enjeu majeur de santé publique. Pour 10 % des patients, elle devient chronique, affectant significativement leur qualité de vie. Bien que des programmes comme la Restauration Fonctionnelle du Rachis (RFR®) incluent des approches éducatives pour déconstruire les croyances erronées, l'efficacité à moyen terme reste incertaine. Cette étude proposerait d'identifier les facteurs déterminants permettant de maintenir les bénéfices acquis post-rééducation.

Méthodologie

Cette étude prospective et multicentrique explorerait les facteurs sociaux et personnels influençant l'efficacité des soins. Des entretiens, des bilans physiques et des auto-questionnaires seront réalisés à plusieurs moments (avant, après, et à 4 et 6 mois post-programme). Les patients seront classés en sous-groupes à l'aide d'outils de Machine Learning (Random Forest, XGBoost), permettant de mieux comprendre leur réponse aux soins.

Discussion

L'analyse mettrait en évidence une hétérogénéité dans la réponse aux traitements, soulignant la nécessité d'une prise en soin individualisée. La stratification des patients en "répondants" et "non-répondants" permettrait de proposer des interventions mieux adaptées.

Conclusion

L'intégration des techniques d'intelligence artificielle dans la personnalisation des soins pourrait améliorer l'efficacité des programmes de rééducation fonctionnelle. Ces outils offriraient la possibilité de mieux cibler les interventions selon les spécificités individuelles des patients atteints de lombalgie chronique, optimisant ainsi les résultats cliniques et thérapeutiques.

Mots-clés : Lombalgie chronique, restauration, rééducation, intelligence artificielle, apprentissage automatique, sociologie, condition physique

Faits saillants

- Première étude associant déterminants sociaux et apprentissage automatique en rééducation.
- Approche novatrice intégrant données physiques, fonctionnelles et sociales dans l'analyse.
- Identification des profils répondeurs/non-répondeurs pour personnaliser les soins en RFR®.
- Suivi longitudinal complet pour évaluer l'efficacité des programmes à moyen terme.

1. Introduction

La lombalgie, douleur localisée au niveau de la région lombaire, est une affection qui concernera environ 83 % des français à un moment de leur vie [1]. 10% de ces cas ont/auront des douleurs qui persistent/persisteront au moins pendant trois mois pouvant aller jusqu'à plusieurs années [2]. Cette affection porte en son histoire et dans l'image qui lui est associée, une connotation et une interprétation négative qui vont de pair aux notions de catastrophisme et d'exacerbation [3]. Le dos, et plus particulièrement la région lombaire, est souvent perçue comme une zone articulaire fragile, qu'il convient de protéger et de ne pas trop solliciter. Ceci amène à des comportements inadaptés lorsque les individus font face à des épisodes douloureux. Le versant cognitivo-comportemental joue donc un rôle majeur dans la durée et la persistance de ces douleurs. Cela inclut toutes les croyances (souvent erronées) et les connaissances (inexactes) acquises au cours de la vie et propre à nos sociétés, mais qui depuis peu évoluent [4]. Cette prise de conscience du grand public rejoint les nombreux modèles biopsychosociaux déjà existants, essayant d'expliquer les mécanismes cognitivo-comportementaux des individus [5], [6], [7].

Ces modèles biopsychosociaux sont des outils pour la recherche mais aussi pour les cliniciens. Les programmes de type Restauration Fonctionnelle du Rachis (RFR®) intègrent ces modèles de manière informelle et formelle par l'intermédiaire de séances d'éducation thérapeutique dont l'objectif est de déconstruire les idées reçues et d'apporter des solutions [8]. Ces programmes visent à rendre les patients autonomes dans la gestion de la douleur en leur fournissant des outils adaptés. Cela comprend des exercices reproductibles chez soi, un apprentissage des bonnes pratiques et de l'auto-gestion de sa santé. Malgré ces approches qui sont de plus en plus proposées dans les programmes, nous constatons une perte d'efficacité à moyen terme pour la plupart des patients. Cette perte d'efficacité, comprise entre 20 et 50%, prend en compte une diminution des gains obtenus que ce soit sur le plan physique mais aussi sur le bien-être général [9]. Mais ceci est sans parler que les études n'ont pas la main sur les consultations post-séjour.

En effet, lors des consultations post-séjour, il existe une déperdition conséquente amenant à requalifier la pertinence des recherches effectuées sur l'efficacité des programmes de reconditionnement à moyen et long terme. Bien souvent, les résultats des tests réalisés ne prennent pas en compte les pertes notables des sujets qui ne se sont pas présentés, et aucune étude à notre connaissance ne mentionnent leurs spécificités et leurs caractéristiques, qu'elles soient générales ou particulières.

La question se pose alors de savoir pourquoi certains patients parviennent à conserver les bénéfices de ces programmes, tandis que d'autres rechutent ? Quels en sont les déterminants et les facteurs explicatifs ? Existe-t-il réellement des facteurs individuels spécifiques ou cela relève-t-il d'une variable aléatoire ?

2. Méthodologie

2.1. Objet d'étude

La littérature scientifique, sans être spécifique à la problématique de la lombalgie, peut nous aider à comprendre et à identifier certains facteurs socio-personnels pouvant influencer l'adhésion à une activité physique, et dans notre cas au programme proposé. Ces études ne s'appliquent pas dans un contexte de rééducation et de restauration fonctionnelle du rachis mais elles sont présentées afin de mieux comprendre nos hypothèses, et ainsi cette recherche.

Tout d'abord, les causes de la lombalgie sont différentes. Certaines sont les résultantes de pathologies rachidiennes alors que d'autres ne s'expliquent pas (non-spécifique) [8]. De la même manière, la relation et l'interprétation de la douleur sont différentes. Nous réagissons tous différemment face aux douleurs. La perception intrinsèque de la douleur, particulièrement dans le contexte de la lombalgie, est un domaine de recherche riche en complexité, intégrant des dimensions biologiques, cognitives, émotionnelles, et socioculturelles.

D'un point de vue cognitif, la douleur lombaire est perçue à travers un prisme de croyances, d'attentes et d'expériences antérieures, qui peuvent influencer non seulement l'intensité de la douleur ressentie, mais aussi la manière dont elle est interprétée par l'individu. Par exemple, la théorie de la « catastrophisation » propose que les patients qui anticipent le pire scénario possible concernant leur douleur ont tendance à ressentir une intensité de douleur accrue et à développer une incapacité fonctionnelle plus sévère [10], [11]. Sur le plan émotionnel, elle est souvent liée à des états d'anxiété et de dépression, qui non seulement exacerbent la douleur perçue, mais contribuent également à un cycle vicieux de détérioration de la qualité de vie [12], [13]. Les études montrent que la co-occurrence de troubles dépressifs et de douleurs lombaires est fréquente, et que la dépression peut non seulement intensifier la douleur, mais aussi prolonger sa durée en inhibant les mécanismes cognitifs et émotionnels de résilience [14], [15].

Sociologiquement, la lombalgie s'accapare de stéréotypes et d'attentes sociales qui influencent à la fois la perception personnelle de la douleur et la réponse sociale à cette douleur. Par exemple, dans de nombreuses sociétés, le mal de dos est souvent minimisé ou perçu comme une condition bénigne, malgré son potentiel d'invalidité sévère, ce qui peut conduire à une stigmatisation des individus souffrant de lombalgie chronique [16], [17]. Cette stigmatisation peut exacerber la souffrance en renforçant un sentiment d'isolement et en diminuant la légitimité perçue de la douleur par l'individu lui-même. Les patients peuvent intérioriser ces stéréotypes, ce qui influence leur propre perception de la gravité de leur condition et peut les amener à minimiser ou à ignorer leur douleur, retardant ainsi les soins nécessaires.

Les recherches en sociologie de la santé ont également mis en évidence que les perceptions de la lombalgie sont influencées par les discours médicaux et médiatiques qui encadrent le mal de dos dans un contexte de productivité et d'aptitude au travail [18]. La douleur est ainsi souvent interprétée non seulement comme une indication de problème de santé, mais aussi comme une menace à la performance professionnelle, ce qui peut intensifier la détresse psychologique associée à la lombalgie et renforcer le cycle de chronicité.

Les parcours de vie et les processus de socialisation, façonnés par de multiples déterminants tels que le milieu socio-économique, le capital éducatif, les expériences familiales et les interactions sociales, jouent un rôle fondamental dans la structuration des comportements, des croyances et des valeurs des individus. Ces dimensions structurantes influencent profondément la manière dont chaque sujet appréhende et gère sa santé et son bien-être. La socialisation, en tant que processus cumulatif d'intériorisation de dispositions sociales, conduit à l'incorporation d'habitus qui orientent de manière durable les pratiques de santé [19]. Par exemple, un individu socialisé dans un contexte valorisant l'activité physique est susceptible de développer une gestion proactive de la santé, incluant de manière routinière la pratique d'activités physiques et sportives. En revanche, lorsque l'individu évolue dans un environnement où l'activité physique est peu valorisée, ses pratiques de santé peuvent s'inscrire dans une logique de passivité, marquée par la sédentarité. Cette différenciation des pratiques s'explique souvent par des inégalités sociales structurelles, où les groupes sociaux disposant de capitaux culturels et économiques limités sont davantage enclins à adopter des comportements sédentaires et à accorder moins d'importance aux soins préventifs [20], [21]. Ces inégalités dans l'accès aux ressources et aux opportunités de pratique physique sont

particulièrement manifestes dès la petite enfance, conditionnant ainsi les trajectoires de santé de manière différenciée en fonction de l'origine sociale [22].

L'analyse des usages sociaux du corps, telle que formulée par Luc Boltanski,²⁰ met en lumière que les pratiques corporelles et sportives ne sont pas uniquement des manifestations individuelles, mais qu'elles relèvent également de constructions sociales et symboliques qui traduisent les positions de classe. Les sujets appartenant aux milieux favorisés tendent à instrumentaliser le sport comme un vecteur de socialisation secondaire, cultivant des dispositions corporelles conformes aux normes dominantes de maîtrise de soi et de discipline. À l'inverse, dans les milieux populaires, les pratiques sportives peuvent être orientées différemment, souvent associées à des logiques de relâchement ou d'expression de soi, plutôt qu'à des impératifs de performance [23].

Les travaux de Christine Mennesson et ses collaborateurs illustrent également que les pratiques physiques et sportives chez les enfants sont socialement différenciées. Dans les groupes sociaux favorisés, ces pratiques sont souvent encouragées dans une optique de socialisation à la performance et au développement de compétences socialement valorisées, renforçant ainsi l'habitus sportif. En revanche, dans les milieux populaires, ces pratiques peuvent répondre à d'autres logiques, davantage orientées vers l'expression de soi et le plaisir immédiat, sans la même pression normative à la performance [23]. Ces différenciations dans les usages des pratiques sportives contribuent à la reproduction des inégalités sociales de santé dès la petite enfance, conditionnant des trajectoires différenciées selon le contexte social [22].

Les transformations sociétales récentes, en particulier en matière de normes sociales et de parcours de vie, ont profondément modifié la manière dont les individus conçoivent et gèrent leur rapport à la santé. Comme le souligne Stefano Cavalli, l'individualisation croissante des trajectoires de vie – caractérisée par une flexibilité et une pluralité de choix – a intensifié la responsabilisation individuelle face à la santé [24]. Dans ce contexte, la carrière d'une personne atteinte de maladie chronique devient une illustration de cette évolution : influencée par les normes de vie et les modèles de socialisation, elle reflète un processus d'adaptation continue et de réévaluation des rôles sociaux, où l'individu navigue entre les attentes culturelles et les réalités de ses limitations physiques [25], [26]. Les parcours de vie, de plus en plus individualisés, incitent ainsi chacun à prendre en charge sa santé non seulement en fonction de critères médicaux, mais aussi en réponse aux attentes sociales et culturelles. Cette responsabilité se manifeste notamment par l'adoption de pratiques physiques, qui ne

répondent plus seulement à un impératif de bien-être, mais deviennent également des marqueurs de gestion proactive et responsable de soi.

La carrière de l'individu, dans ce cadre, se construit par phases de rupture et de reconstruction, redéfinissant les comportements et l'identité à chaque étape. Contrairement à une simple succession de symptômes physiques, cette carrière incarne un véritable parcours social : l'individu doit alors non seulement gérer la chronicité de la maladie, mais aussi réconcilier ses aspirations personnelles avec les attentes extérieures. Ce processus d'ajustement reflète une trajectoire dynamique, où chaque expérience de la maladie modifie la manière dont la personne perçoit ses capacités et son rôle dans la société, redéfinissant en continu son rapport à la santé [26], [27].

En première instance, l'apparition de la maladie constitue souvent une rupture significative par rapport à la « normalité », créant une discontinuité dans la trajectoire de vie qui entraîne la réévaluation de rôles sociaux, comme ceux d'employé, de partenaire, ou de parent. Par exemple, la douleur et les limitations physiques associées à la lombalgie chronique remettent en question la capacité d'un individu à maintenir un emploi, à assumer ses responsabilités familiales, ou à participer aux activités sociales, influençant ainsi l'auto-perception et la perception par autrui de ses capacités. Cela amène souvent les patients à adopter des stratégies d'adaptation qui se manifestent par des comportements de retrait ou, au contraire, par une lutte pour maintenir leur statut antérieur malgré les obstacles [27].

Cette carrière est également marquée par un processus d'interaction avec les autres, où l'individu adapte et réinterprète sa place en société. L'approche interactionniste de la carrière par Darmon révèle comment les échanges avec les soignants [25], les proches, et les pairs jouent un rôle déterminant dans la manière dont les personnes atteintes de lombalgie chronique façonnent leur identité et leur rapport à la douleur. Ces interactions ne sont pas seulement des moments de partage d'information ou de soutien, mais des contextes dans lesquels l'individu découvre de nouvelles façons d'habiter son corps et de réagir aux normes sociales. Ces dynamiques interactionnelles influencent directement les comportements liés à la gestion de la maladie : l'individu pourrait ainsi adopter des attitudes d'hypervigilance ou, au contraire, développer une tolérance accrue à la douleur, en réponse aux attentes implicites ou explicites de son entourage.

L'évolution de cette carrière influence profondément sur la qualité et la pertinence des prises en soin (PES). Comprendre le parcours de vie des personnes atteintes de lombalgie chronique,

leurs étapes de socialisation, et les rôles qu'elles sont prêtes ou non à endosser, permet aux professionnels de proposer des interventions qui tiennent compte des besoins psychosociaux du patient, et pas uniquement des symptômes physiques.

Claire Perrin, dans son étude sur "Analyse des relations entre le rapport aux APS (Activités Physiques et Sportives) et les conceptions de la santé", met en évidence que les conceptions de la santé sont étroitement liées aux représentations et aux usages sociaux de l'activité physique. Selon elle, l'activité physique est souvent perçue à travers une double fonction : elle est à la fois un outil de maintien de la santé et un moyen de correspondre aux normes sociales de performance et de contrôle du corps [28]. Ces pratiques s'inscrivent donc dans un cadre normatif, où la santé est envisagée non seulement comme un état, mais comme un ensemble de pratiques à réaliser activement. En ce sens, l'adoption d'activités physiques peut être vue comme un mode d'engagement vis-à-vis des normes dominantes, valorisant l'efficacité, la discipline et la maîtrise corporelle.

Dans le cas particulier des individus atteints de maladies chroniques, les enjeux normatifs autour de l'activité physique prennent une dimension encore plus prononcée. Perrin et Mino [29], en s'appuyant sur la philosophie de Georges Canguilhem, soulignent que ces personnes sont souvent amenées à redéfinir leur rapport à l'activité physique dans le cadre de l'adaptation à des normes corporelles fluctuantes. L'Activité Physique Adaptée (APA) devient ainsi un espace de négociation entre les exigences médicales et les normes sociales qui valorisent la performance et l'optimisation du corps, même en contexte de maladie [29]. Cette tension entre la normalisation et l'adaptation est au cœur des trajectoires de santé des personnes vivant avec des maladies chroniques, où le corps malade doit continuellement s'ajuster aux normes sociales de la santé et de l'efficacité corporelle.

Dans le cas de la lombalgie, les différences de socialisation, de parcours de vie et de contexte social jouent un rôle déterminant dans la manière dont les individus gèrent leurs symptômes. Une personne ayant intégré l'activité physique de manière régulière, dans un environnement social qui valorise ces pratiques, est souvent mieux préparée à adopter des comportements favorables à la gestion de la lombalgie. À l'inverse, un individu moins actif et ayant une perception négative de l'exercice physique peut être enclin à éviter toute forme d'activité, ce qui peut aggraver la douleur et favoriser la chronicité du problème. C'est précisément pour répondre à cette problématique que les programmes de réadaptation fonctionnelle, tels que les programmes RFR®, visent à fournir aux patients des outils adaptés, indépendamment de leur expérience antérieure avec l'activité physique.

Toutefois, les déterminants sociaux et les parcours de vie des patients sont encore trop souvent sous-estimés dans l'élaboration des programmes de PES. Or, ces éléments peuvent exercer une influence directe sur l'efficacité des interventions thérapeutiques. Il apparaît donc nécessaire de prendre en compte non seulement l'état actuel du patient, mais aussi son contexte social et ses antécédents en matière d'activité physique, afin d'adapter au mieux les soins.

En nous appuyant sur les éléments discutés précédemment, nous émettons l'hypothèse qu'il existe des profils distincts de patients dans les programmes de PES: les patients « répondeurs », qui profitent pleinement des interventions proposées, et les patients « non-répondeurs », pour lesquels les résultats sont moins favorables. Cette différenciation pourrait s'expliquer par l'ensemble des points énoncés jusqu'à présent. Pour les patients « non-répondeurs », il serait nécessaire d'orienter et d'adapter la PES de manière plus ciblée et personnalisée. Cette personnalisation des soins pourrait permettre de mieux répondre aux besoins spécifiques de ces patients et d'améliorer l'efficacité des programmes.

Il est question dans ce travail d'identifier les déterminants sociaux et personnels influençant la qualité et l'efficacité de la PES à court et moyen terme (environ 6 mois).

2.2. Hypothèses

- Divers déterminants sociaux influencent l'efficacité de la prise en soin et par conséquent, la caractérisation sociale des patients permet d'individualiser le programme, le rendant plus efficace et adapté.
- La caractérisation sociale permet d'orienter les patients vers une prise en soin adaptée, individualisée et appropriée.

2.3. Conception de l'étude

Cette étude randomisée et prospective sera multicentrique. Elle sera réalisée au sein de structures rééducatives. Le nombre de sujet requis pour cette étude ne peut être défini pas un G-power étant donné qu'elle est multidimensionnelle. Nous opterons donc pour une récolte de sujets jusqu'à saturation des données.

L'objectif principal de cette étude est de caractériser les patients atteints de lombalgie en explorant divers aspects de leur vie tout en les associant aux bénéfices obtenus sur leur condition physique, le versant cognitif et leur qualité de vie . Pour ce faire, nous mettons en place un protocole de recherche structuré autour d'entretiens directifs et de passation de bilans physiques et d'auto-questionnaires.

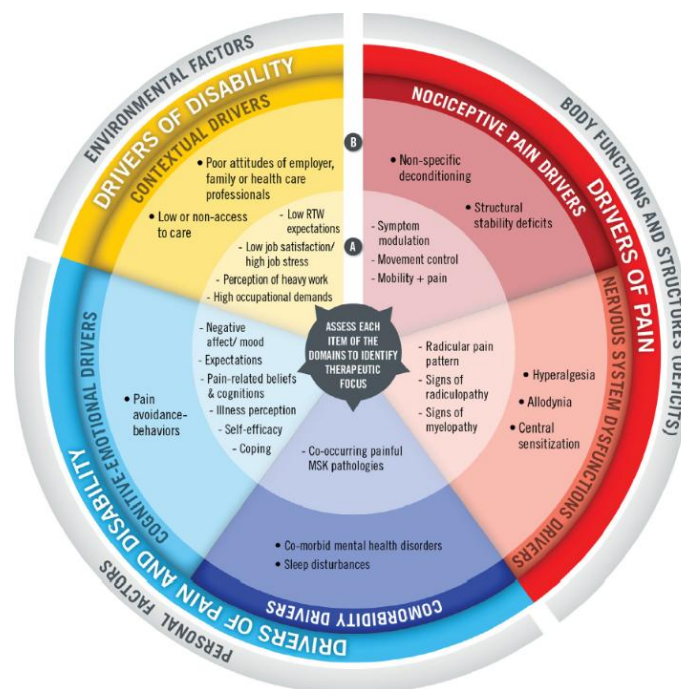
2.3.1. Cadre théorique de la stratification et de l'individualisation

Comme mentionné dans l'introduction, chaque cas de lombalgie est unique, et ce pour de multiples raisons. Comment envisager qu'un programme standardisé tel que le RFR® puisse réellement convenir à chaque patient sans prendre en compte son parcours de vie et ses aspects cognitifs et comportementaux ? Cette question soulève des enjeux cruciaux en matière de pertinence et d'efficacité des soins proposés dans les structures de santé. L'objectif fondamental est que chaque patient, quelle que soit sa pathologie, ait accès aux soins les plus adaptés, efficaces et efficaces possible. Ce protocole de recherche vise à proposer une solution intégrative et écologique, en s'appuyant sur des axes sociologiques, physiologiques et psycho-comportementaux. Cette solution serait permise par la mise en place de méthodes de classification et de clustering. L'objectif final est de filtrer et d'orienter les patients des soins adaptés.

La littérature scientifique propose diverses méthodes de classification et d'adaptation des soins. Depuis les années 1980, des stratégies de classification se sont multipliées : certaines s'attachent aux mécanismes physiopathologiques sous-jacents, d'autres aux réponses individuelles au traitement, tandis que d'autres stratifient selon le niveau de risque.

Le MGVDI (Modèle de Gestion des Vecteurs de Douleur et d'Incapacité), présenté ci-dessous, est un outil qui illustre une première méthode [30].

Figure 1 Modèle de Gestion des Vecteurs de Douleur et d'Incapacité



Ce modèle, fondé sur une batterie de tests exhaustive, est conçu pour orienter le professionnel vers des soins adaptés aux causes potentielles de la douleur. Toutefois, bien que pertinent pour des praticiens spécialisés dans la lombalgie chronique, il omet de prendre en compte le contexte sociologique et l'expérience du patient. Il ne fournit pas non plus de recommandations sur les modalités de traitement (fréquence, intensité). De plus, à notre connaissance, aucune étude n'a quantifié sa capacité prédictive à rendre un soin plus adapté.

D'autres outils de classification existent également, tels que la méthode McKenzie et le questionnaire STarT Back Tool, qui aident à déterminer si ces mêmes méthodes spécifiques sont applicables et pertinentes en fonction des profils des patients. Mais ces méthodes, bien qu'utiles, présentent des limites dans leur application clinique, en particulier en ce qui concerne la fiabilité, la précision de la classification et l'orientation très simplifiée du traitement [31], [32], [33].

D'autres études ont mis en place des techniques de clustering [34], [35], c'est-à-dire des méthodes regroupant des relations et des interactions similaires entre plusieurs jeux de données et variables. Le clustering est une technique d'apprentissage qui peut être supervisée ou non-supervisée. Souvent associées au Machine Learning (ML) ou au Deep Learning (DL), elles sont de plus en plus utilisées dans le domaine de la santé et de la recherche. C'est notamment le cas de la lombalgie chronique qui représente un domaine de recherche en pleine expansion (bien que cette application reste encore à ses débuts). Les avancées technologiques récentes, en particulier dans l'analyse de grandes bases de données cliniques, offrent un potentiel considérable pour la classification des patients et la prédiction de l'efficacité des traitements [36], [37], [38].

Le ML, avec sa capacité à analyser des ensembles de données de grande taille et à détecter des relations non linéaires, permet d'explorer des données multidimensionnelles. Ces algorithmes apprennent à partir de données sans qu'il soit nécessaire de programmer explicitement toutes les règles d'interprétation. En pratique, cela se traduit par une capacité à automatiser l'identification des patients présentant des facteurs spécifiques influençant la réponse aux soins. Malgré cette promesse, la majorité des études se sont concentrées jusqu'à présent sur des classifications binaires, telles que la présence ou l'absence de lombalgie, sans aller au-delà dans la stratification des sous-groupes cliniques en fonction des résultats des traitements. La littérature montre que seulement une poignée d'études s'est intéressée à la prédiction de la réponse aux traitements spécifiques à travers des algorithmes d'IA, ce qui

limite encore l'impact pratique de ces approches dans la PES des patients atteints de lombalgie chronique [31], [39].

En effet, toutes les méthodes de classification exposées présentent plusieurs limites : elles négligent souvent les facteurs psychosociaux et démontrent une faible capacité prédictive au niveau clinique, en particulier pour les patients atteints de lombalgie chronique [32]. De plus, des études et méta-analyses offrent des résultats contradictoires quant à leur efficacité [33], [40], [41]. Par conséquent, ces approches ne permettent pas de saisir pleinement l'hétérogénéité des patients souffrant de lombalgie non spécifique [42]. Ces constatations justifient la recherche de méthodes de classification plus complètes et adaptées aux caractéristiques complexes de cette population.

Pour que ces outils basés sur l'IA atteignent un niveau d'efficacité comparable ou supérieur aux méthodes traditionnelles, il est impératif de s'appuyer sur des ensembles de données plus vastes, avec des méthodes de ML et de clustering adaptées et des bilans complets afin d'identifier tous les déterminants pouvant expliquer un niveau d'efficacité à court et moyen terme [31], [39], [43].

La force de nos recherches résulte sur les points précédemment énoncés. Il est question de rendre compte de l'efficacité et de l'efficience des différents types de soins que nous proposerons avec un regard multidimensionnel et complet. L'objectif est de faire ressortir des déterminants influençant les gains obtenus en fonction des types de soins auxquels les sujets seront soumis. Cela dégagera des profils répondants et non-répondants qui, à terme, permettra d'orienter efficacement les patients vers des modalités de soins adaptés.

2.3.2. Critères d'inclusion et d'exclusion des sujets

Pour participer à cette étude, les critères d'inclusion requièrent que les patients aient une lombalgie depuis plus de trois mois et aient connu un échec thérapeutique. De plus, pour les participants âgés de plus de 40 ans, un test d'effort sera nécessaire pour évaluer leur aptitude cardio-respiratoire. À l'inverse, certains critères d'exclusion seront établis pour garantir l'homogénéité des groupes étudiés : les patients qui auront subi une arthrolyse poly-étagée et ceux âgés de plus de 65 ans seront exclus. Seront également exclus les patients ayant subi une chirurgie discale dans les six mois précédents, ceux avec une prothèse totale de genou ou de hanche, un indice de masse corporelle supérieur à 30, ou ceux ayant déjà suivi un programme spécifique pour le dos.

Ces critères relèvent de la pertinence des soins proposés surtout lorsqu'il s'agit de programme de type RFR®. Ces critères ne sont que très peu contraignant dans le cadre de notre recherche et dans son recrutement. Les groupes représenteront la grande majorité des cas de lombalgie présents dans les structures de santé. Aucune méthode de randomisation n'a été déployée et ce pour les raisons détaillées précédemment.

2.3.3. Interventions

Trois types de programmes seront proposés. Ils intégreront des modes de PES très largement étudiés dans la littérature. Le Programme Autonomie « PA » a pour vocation à rendre la patient acteur de sa santé par l'autonomie qui lui est permise (muscultation et travail cardio-vasculaire) et par l'apport de séances d'éducation thérapeutique offrant des outils adaptés. Le Programme Reconditionnement « PR » n'a que par intérêt de reconditionner le patient sans se soucier de l'après-PES. Il est question ici de suivre des séances de manière très structurée. Le Programme Léger « PL » a quant à lui une forme différente des deux programmes. Il reprend les fondamentaux du PA mais moins intensif et plus étalé dans le temps. Nous exposons les modalités des 3 programmes dans le tableau 1 ci-dessous :

Tableau 1 Caractéristiques des programmes étudiés

	PA	PR	PL
Nombre de semaines	4	4	12
Nombre d'heures de séances encadrées (et totales)	83 (130)	116 (116)	8 (24)
Nombre d'heures totales encadrées par semaine	21	29	2
Types de séances	Renforcement musculaire sur des machines guidées et sur tapis (autonomie)	Renforcement musculaire sur des machines guidées et sur tapis	Renforcement musculaire sur des machines guidées et sur tapis (autonomie)
	Travail cardio-vasculaire (autonomie)	Travail cardio-vasculaire	Travail cardio-vasculaire (autonomie)
	Kinésithérapie	Kinésithérapie	X

	Ergothérapie	Ergothérapie	X
	Activité Physique Adaptée	X	X
	Education thérapeutique	X	Education thérapeutique
	Autres corps de métiers (psychologue, nutritionniste etc.)	Autres corps de métiers (psychologue, nutritionniste etc.)	X
Objectifs principaux	Reconditionnement général avec la volonté forte de rendre autonomie le patient.	Reconditionnement général sans pour autant avoir de séances d'éducation proposées.	Reconditionnement général avec une moindre intensité et la volonté de rendre autonome le patient. L'objectif est que ces exercices soient reproduits régulièrement en dehors des heures en centre.

2.3.4. Paramètres évalués

2.3.4.1. Mesure sociologique : passation d'entretiens directs

Ces entretiens, d'une durée de trente minutes à une heure, sont conçus pour recueillir des informations qualitatives détaillées sur plusieurs dimensions de la vie ainsi que sur la conception et le rapport aux activités physiques et à la santé des patients:

- Contexte de vie pendant la jeunesse/socialisation primaire : exploration des habitudes familiales, des influences sociales, et des premières expériences liées à la santé et à l'activité physique.

- Évolution de la pratique d'activité physique : analyse des changements dans la pratique sportive et l'activité physique tout au long de la vie, incluant les périodes d'inactivité ou de reprise.
- Évolution de la lombalgie : évaluation des premières apparitions de la douleur, des variations, de son intensité, de sa fréquence et de son impact sur la qualité de vie.
- Comportements face à la douleur : identification des stratégies de gestion de la douleur adoptées par les patients, qu'il s'agisse de recours aux soins médicaux, d'automédication ou d'autres approches thérapeutiques.
- Rapport aux soins et hygiène de vie : exploration des attitudes des patients vis-à-vis des soins de santé, y compris l'adhésion aux recommandations médicales, et leur rapport à l'hygiène de vie tels que l'alimentation et la gestion du stress.
- Événements marquants de la vie : recueil des événements significatifs ayant potentiellement influencés leur santé physique et mentale, comme les traumatismes, les pertes d'un proche ou les expériences de maladie.

Ces entretiens permettent de recueillir des données fournissant une vue d'ensemble des comportements de santé et des perceptions de la douleur des patients.

2.3.4.2. Mesures physiques et de perceptions fonctionnelles et cognitives (PROMs)

En parallèle des entretiens, des données physiques sont collectées lors de bilans réalisés à plusieurs moments: à l'entrée du programme de réadaptation, à la sortie, puis à 4 et 6 mois après le séjour. Ces bilans incluent des mesures de flexibilité du tronc et des membres inférieurs, de force et d'endurance musculaire.

Les participants complètent également des auto-questionnaires à ces différents moments, qui visent à recueillir des informations sur leur état de santé perçue, leur qualité de vie, leur niveau d'activité physique, et leur état psychologique. Ces données sont cruciales pour comprendre l'évolution des patients tout au long du programme et au-delà. Ils sont considérés comme des outils PROMs (Patient Reported Outcome Measures) [44]. Ils intègrent la perception et l'efficacité subjective des soins apportés aux patients.

Le tableau n°2 est un récapitulatif des bilans physiques et auto-questionnaires réalisés aux différents moments de sa PES :

Tableau 2 Protocole de tests et de mesures

		Début	Fin	4 mois	6 mois
Informations diverses	Données anthropométriques, situation professionnelle etc.				
Parcours de vie	Entretien				
Mobilité	Distance doigts-sol [45]				
	Schober				
	Inclinaison (droite/ gauche)				
	Rotation (droite/ gauche)				
	Psoas (droite/ gauche)				
	Ischio-jambiers (droite/ gauche)				
Force musculaire	Test isocinétique (30 et 60°/s)				
Endurance musculaire	Ito-shirado [46]				
	Sorensen [47]				
	Gainage latéral (Droite/gauche)				
	Killy				
Auto-questionnaires	ONAPS-PAQ [48]				
	DALLAS [49]				
	EIFEL [50]				
	GHQ-12 [51]				
	FABQ [52]				

Zone verte : moment où les mesures sont faites

Zone blanche : moment où les mesures ne sont pas faites

2.3.4.3. Critères de jugement principaux

Nous allons dans un premier temps avoir une approche de régression c'est-à-dire que nous analyserons l'évolution des scores obtenus aux différents tests. Nous en estimerons également les valeurs dans le cas de nos modèles prédictifs.

Nous nous intéresseront également à certains paramètres que nous jugeons comme des critères de réussite d'un programmes :

- Une augmentation des activités physiques personnelles

- Une réduction de la douleur
- Un cadre professionnel optimisé (adaptation de poste, changement de métier, réinsertion professionnelle)

2.3.5. Analyses et méthodes statistiques

2.3.5.1. Points de vigilance

L'utilisation de méthodes de ML dans la recherche adressée à la thématique de la douleur, notamment des techniques telles que UMAP, les Random Forest ou bien XGBoost, demeure relativement rare, surtout dans des contextes de faible effectif. Les algorithmes comme Random Forest et les modèles de clustering non supervisés, tout en étant puissants pour identifier des sous-groupes cliniquement pertinents, nécessitent souvent des ensembles de données volumineux pour atteindre des niveaux de précision statistique satisfaisants. Cependant, nous mettons le point sur les caractéristiques de nos sujets. Si elles reflètent les cohortes déjà existantes dans d'autres recherches, nous pourrions nous satisfaire d'avantage des performances que nous obtiendrons.

2.3.5.2. Préparation des données

Les données collectées incluent des mesures variées, allant de caractéristiques physiologiques, comportementales, fonctionnelles, psychologiques et sociales. Ces variables peuvent se trouver sur des échelles très différentes, ce qui peut fausser les analyses si elles ne sont pas standardisées. Afin d'éviter ce déséquilibre et de garantir que toutes les variables contribuent équitablement à l'analyse, nous normaliserons les données. Cette transformation aligne les variables sur une échelle commune (de 0 à 1).

$$x^{normalisé} = (x - x^{min}) / (x^{max} - x^{min})$$

Or certaines autres fonctions et algorithmes que nous allons développer nécessitent des données standardisées. La standardisation est une technique de prétraitement qui centre et réduit les variables en leur soustrayant leur moyenne (μ) et en divisant par leur écart-type (σ) : $x^{standardisé} = (x - \mu) / \sigma$. Cette transformation permet d'obtenir des variables de moyenne nulle et d'écart-type unitaire, facilitant ainsi la comparaison entre des variables d'échelles différentes et optimisant les performances de nombreux algorithmes statistiques.

$$x^{standardisé} = (x - \mu) / \sigma$$

Nous appliquerons des techniques d'imputation adaptées aux valeurs manquantes. Le test de Little a révélé que nos données ne sont pas MCAR, mais probablement MAR ou MNAR, nécessitant une approche spécifique. Si elles ne sont pas MCAR, nous utiliserons MICE (Multiple Imputation by Chained Equations), un processus itératif de régression prenant en compte l'incertitude et préservant les relations entre variables [53]. Nous l'avons adapté avec une régression Bayésienne (Bayesian Ridge) pour assurer la cohérence clinique. En cas de données MCAR, nous privilégierons des méthodes plus simples comme la moyenne ou la médiane.

2.3.5.3. Exploration et réduction de dimensionnalité : approche multi-méthodes

Pour commencer, nous allons devoir réduire les dimensions au vue du nombre de variables conséquent. Pour cela nous allons déployer 3 techniques bien distinctes afin de capter le plus d'information possible en fonction de la nature de nos données. L'Analyse en Composantes Principales (ACP) sera notre première technique qui visera à transformer des variables corrélées en nouvelles variables non corrélées, appelées composantes principales [54]. L'ACP opère en construisant des combinaisons linéaires des variables originales, ordonnées selon leur importance dans l'explication de la variance totale des données.

Pour traiter les variables qualitatives, l'Analyse des Correspondances Multiples (ACM) étend les principes de l'ACP aux données catégorielles [55]. Cela nous permettra d'ajouter nos variables sociologiques en vue d'explorer prématurément les premières interactions multivariées. Les taux d'inertie sont ajustés selon la méthode de Benzécri pour fournir une estimation plus précise de l'information effectivement extraite des données.

Nous déploierons pour terminer une analyse UMAP (Uniform Manifold Approximation and Projection) et T-SNE (t-Distributed Stochastic Neighbor Embedding) afin de compléter notre approche préliminaire [56-57]. Ce modèle aborde une approche non-linéaire rendant ainsi son interprétabilité plus difficile mais plus robuste.

2.3.5.4. Classification Ascendante Hiérarchique : Structuration des profils patients

L'objectif de cette étape sera d'identifier des groupes homogènes de patients en fonction de leurs caractéristiques physiologiques, psychologiques et sociologiques afin de mieux comprendre les profils ayant des trajectoires différenciées face aux traitements de réhabilitation. Nous précisons qu'il s'agit d'une première analyse exploratoire basée sur nos méthodes de réduction de dimension précédemment détaillées. L'ensemble de nos données qualitatives seront factorisées et encodées via Ordinal Encoder de Scikit-learn. Ces données seront ensuite standardisées.

La Classification Ascendante Hiérarchique (CAH) sera utilisée comme première approche exploratoire [58]. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) sera également appliqué afin d'identifier les regroupements denses tout en détectant les observations isolées considérées comme du bruit [59]. L'optimisation de ses paramètres ϵ et minPts est réalisée en s'appuyant sur l'analyse des distances k-plus proches voisins, ce qui permet d'ajuster le seuil de densité et d'améliorer la segmentation. En complément, HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) adaptera dynamiquement la détection des clusters en ajustant la granularité en fonction de la densité locale, optimisant ainsi la robustesse des groupes obtenus [60].

Le clustering spectral ainsi que le modèle de mélanges gaussiens (GMM) seront également mobilisés afin d'attribuer aux patients une appartenance probabiliste à un ou plusieurs clusters, offrant une flexibilité supplémentaire dans la segmentation des profils [61]. Pour GMM, le choix optimal du nombre de clusters reposera sur les critères AIC (Akaike Information Criterion) et BIC (Bayesian Information Criterion), qui permettront d'optimiser la complexité du modèle tout en assurant une bonne généralisation des regroupements [62-63].

Une agrégation ensembliste des résultats obtenus via nos techniques de clustering sera mise en œuvre. Cette approche visera à stabiliser la classification en s'appuyant sur plusieurs modèles complémentaires afin d'identifier des profils robustes.

L'évaluation des clusters reposera sur l'Adjusted Rand Index (ARI) ainsi que la silhouette est afin d'évaluer la qualité des clusters en termes de cohésion et de séparation. Une analyse complémentaire des variables discriminantes est réalisée à l'aide de tests statistiques adaptés, en fonction des distributions observées (suivant la logique citée dans la partie précédente).

Cette combinaison méthodologique assure une segmentation optimisée de nos sujets et permettra d'identifier des sous-groupes différenciés.

2.3.5.5. Modélisation prédictive : Approche progressive

Nous commencerons cette étape par une régression logistique qui permettra d'identifier les variables significativement associées aux outcomes (sorties) et d'établir un premier niveau de prédiction.

Les méthodes de régression pénalisée constituent la seconde étape de notre analyse, avec l'application du modèle Elastic Net [64]. Cette technique permettra une sélection des variables pertinentes tout en gérant efficacement la multi colinéarité présente dans les données.

Les méthodes ensemblistes avancées, Random Forest et XGBoost, compléteront notre pipeline [65-66]. Nous implémenteront les variables détectées par nos modèles de régression.

2.3.5.6. Validation et évaluation des modèles

Notre approche de validation et d'évaluation reposera sur la validation croisée répétée (RepeatedKfold). Cette méthode divise les données en 5 sous-ensembles et répète le processus deux fois, assurant ainsi une estimation robuste des performances des modèles. La métrique principale d'évaluation est la RMSE (Root Mean Square Error), complétée par le coefficient de détermination (R^2) pour une vision globale de la qualité des prédictions.

L'optimisation des hyperparamètres sera réalisée via RandomizedSearchCV, qui explore efficacement l'espace des paramètres possibles en utilisant des distributions adaptées à chaque type d'hyperparamètre. Cette approche est plus efficace qu'une recherche exhaustive tout en maintenant une qualité de résultats comparable.

L'évaluation qualitative s'appuiera sur l'analyse SHAP (SHapley Additive exPlanations), qui fournit une interprétation détaillée des prédictions du modèle [67]. Cette analyse est particulièrement pertinente pour nos modèles à base d'arbres (RandomForest et XGBoost), permettant de comprendre l'impact de chaque variable sur les prédictions finales.

3. Discussion

L'objectif global de notre étude est d'identifier un ensemble de facteurs influençant positivement et négativement l'efficacité de trois types de programmes proposés au sein des structures de santé. Les modalités de ces derniers, comme dit dans le cadre théorique, représentent en grande partie les programmes présents dans la littérature et *in fine* dans les

structures de santé en France. Elles se différencient par leurs intensités - 4 et 12 semaines - et par leurs objectifs fondamentaux axés sur le reconditionnement physique ou sur le changement de comportement. La réponse aux soins sera mesurée par une batterie de tests multidimensionnels (physique, qualité de vie, fonctionnel, sociologique) réalisée à différents moments de la PES (au début, à la fin et à 4 et 6 mois).

Avec une récolte de données qui sera conséquente et comprise dans un jeu de variable multidimensionnel, des techniques de clustering seront utilisées et déployées pour l'analyse des résultats et la stratification de « phénotypes ». Comme nous l'avons vu, les études mettant en œuvre des techniques de clustering et de ML n'ont pas intégrées un regard complet explicatif des maintiens des gains post-PES. Aucune étude ne s'est par ailleurs intéressée spécifiquement aux programmes de restauration fonctionnelle du rachis.

La différenciation entre les patients "répondants" et "non-répondants" à un programme de soins soulève la question des déterminants spécifiques. Cette stratification amène à identifier les profils potentiellement répondants aux soins. Il sera donc essentiel de proposer des solutions alternatives pour les patients non-répondants. Cette analyse aura pour objectif d'aider les thérapeutes à prendre des décisions de traitement véritablement adapté au profil de chaque patient. De plus, cette recherche centrée sur diverses approches de soins, offrira des options potentielles pour ajuster ou réorienter les traitements si nécessaire.

L'intégration des technologies de stratification en santé, notamment l'IA et le ML, s'inscrit dans un cadre plus large de médecine et de rééducation personnalisée. Dans la PES des maladies chroniques invalidantes, comme la lombalgie chronique, l'identification de sous-groupes de patients via des algorithmes tels que le Random forest ou XGBoost permettent de mieux comprendre la diversité des réponses aux traitements. En effet, la stratification des patients reposant sur l'analyse de variables multidimensionnelles permet de proposer des soins ciblés et adaptés aux besoins spécifiques de chaque profil. Cette approche pourrait être enrichie par le développement d'outils d'interprétation clinique, permettant aux praticiens de comprendre les facteurs influençant chaque prédiction. L'utilisation d'analyses explicatives, comme les valeurs SHAP, offrirait une lecture claire des variables déterminantes pour chaque patient, facilitant ainsi la personnalisation des interventions thérapeutiques.

Cette approche est particulièrement bénéfique dans les contextes de maladies chroniques où l'hétérogénéité des patients peut compliquer les interventions standardisées. A l'aide du ML, il sera possible de prédire quels patients répondront mieux à un programme d'exercices intensifs

ou à un traitement annexe, réduisant ainsi les rechutes et optimisant les ressources médicales et économiques de tous les acteurs. Ces prédictions, accompagnées d'indicateurs de confiance et d'explications personnalisées, permettraient aux cliniciens d'affiner leurs décisions thérapeutiques tout en conservant leur autonomie de jugement.

Ainsi, la combinaison de la personnalisation des soins et des modèles prédictifs en IA s'annonce comme une voie incontournable pour optimiser l'efficacité des interventions thérapeutiques dans les maladies chroniques.

Bien que les technologies soient prometteuses, il est important de souligner que ces outils ne remplaceront pas les professionnels de santé. En effet, l'IA est avant tout un outil complémentaire qui aide à analyser de grandes quantités de données, à identifier des schémas cachés ou à prédire les réponses aux traitements. Cependant, la prise de décision finale, la compréhension nuancée des besoins des patients, et l'accompagnement humain restent des éléments clés qui nécessitent l'expertise et le jugement des soignants.

Elles ne remplacent pas l'aspect relationnel et clinique du soin, où l'empathie, l'échange et l'expérience professionnelle sont primordiaux. Les soignants, qu'il s'agisse de médecins, kinésithérapeutes, Enseignants en Activités Physiques Adaptées et autres, jouent un rôle central dans l'interprétation des résultats fournis par ces technologies et dans l'adaptation des soins aux contextes individuels des patients. Ainsi, l'IA doit être perçue comme une aide à la décision, non comme un substitut à l'intervention humaine.

En somme, ces outils apportent une plus-value indéniable, mais ils s'inscrivent dans une démarche collaborative où la technologie soutient mais ne remplace pas.

4. Biais potentiels

L'évaluation de l'efficacité des interventions thérapeutiques à court, moyen et long terme se heurte à un défi méthodologique majeur : l'attrition des participants au cours du suivi. Malgré l'obtention du consentement éclairé et l'information complète des participants sur le protocole de recherche, le taux de participation aux bilans post-programme reste un paramètre difficilement contrôlable. Cette problématique soulève non seulement des questions de puissance statistique, mais introduit également un biais de sélection potentiel dans l'analyse des résultats. En effet, les caractéristiques des patients qui se présentent aux suivis post-intervention pourraient différer significativement de celles des non-répondants. Cette auto-

sélection pourrait biaiser l'évaluation de l'efficacité du programme, soit en sur-représentant les patients ayant connu une évolution favorable et désireux de partager leur succès, soit au contraire, en attirant principalement les patients toujours en difficulté et en recherche de solutions. Ce biais d'attrition différentielle constitue donc un enjeu majeur dans l'interprétation des résultats et la généralisation des conclusions de l'étude. Cette recherche offre l'opportunité d'analyser spécifiquement les caractéristiques et les profils des patients présents aux bilans post-programme, permettant ainsi d'identifier les déterminants potentiels de l'adhésion au suivi et de mieux comprendre les facteurs influençant la participation à long terme.

Remerciements

Je tiens à remercier l'Association Nationale de la Recherche et de la Technologie (ANRT) pour le soutien financier accordé dans le cadre de la convention CIFRE, permettant la mise en place de cette étude.

Je souhaite également exprimer ma gratitude envers mes encadrantes pour leur orientation et leurs précieux conseils tout au long de la préparation de ce projet. Un remerciement particulier aux structures de rééducation et aux professionnels de santé qui ont accepté de participer à cette recherche.

Enfin, je remercie par avance les participants pour leur future implication, ainsi que mes collègues pour leur soutien continu dans la réalisation de cette étude.

Bibliographie

- [1] Assurance Maladie. « Mal de dos ? Le bon traitement, c'est le mouvement » : bilan après 2 ans de campagne d'information. <https://www.assurance-maladie.ameli.fr/presse/2019-09-20-cp-campagne-mal-de-dos>. Published September 20, 2019. Accessed October 21, 2024.
- [2] Haute Autorité de Santé. *Lombalgie commune chronique*. HAS; 2024. https://www.has-sante.fr/upload/docs/application/pdf/2024-04/synthese_aps_lombalgie_commune_chronique.pdf.
- [3] Vlaeyen JWS, Kole-Snijders AMJ, Boeren RGB, Van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain*. 1995;62(3):363-372. doi:10.1016/0304-3959(94)00279-n
- [4] Lataste-Munter O, Samandari A, Genevay S. Lombalgies communes chroniques et croyances erronées. *Revue Medicale Suisse*. <https://www.revmed.ch/revue-medicale-suisse/2021/revue-medicale-suisse-760/lombalgies-communes-chroniques-et-croyances-erronees>. Published November 24, 2021.
- [5] Lethem J, Slade PD, Troup JDG, Bentley G. Outline of a fear-avoidance model of exaggerated pain perception—I. *Behaviour Research and Therapy*. 1983;21(4):401-408. doi:10.1016/0005-7967(83)90009-8
- [6] Vlaeyen JWS, Kole-Snijders AMJ, Boeren RGB, Van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain*. 1995;62(3):363-372. doi:10.1016/0304-3959(94)00279-n
- [7] Mayer TG, Gatchel RJ. Functional restoration for spinal disorders : the sports medicine approach. <https://www.ncbi.nlm.nih.gov/nlmcatalog/8711087>. Published 1988.
- [8] Haute Autorité de Santé. Prise en charge du patient présentant une lombalgie commune. Haute Autorité De Santé. https://www.has-sante.fr/jcms/c_2961499/fr/prise-en-charge-du

patient-presentant-une-lombalgie-commune. Published April 4, 2019. Accessed October 21, 2024.

[9] Airaksinen O, Brox JJ, Cedraschi C, et al. Chapter 4 European guidelines for the management of chronic nonspecific low back pain. *European Spine Journal*. 2006;15(S2):s192-s300. doi:10.1007/s00586-006-1072-1

[10] Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: Development and validation. *Psychological Assessment*. 1995;7(4):524-532. doi:10.1037/1040-3590.7.4.524

[11] Turk DC, Okifuji A. Psychological factors in chronic pain: Evolution and revolution. *Journal of Consulting and Clinical Psychology*. 2002;70(3):678-690. doi:10.1037/0022-006x.70.3.678

[12] Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. *Pain*. 1992;50(2):133-149. doi:10.1016/0304-3959(92)90154-4

[13] Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity. *Archives of Internal Medicine*. 2003;163(20):2433. doi:10.1001/archinte.163.20.2433

[14] Gatchel RJ, Peng YB, Peters ML, Fuchs PN, Turk DC. The biopsychosocial approach to chronic pain: Scientific advances and future directions. *Psychological Bulletin*. 2007;133(4):581-624. doi:10.1037/0033-2909.133.4.581

[15] Linton SJ. A review of psychological risk factors in back and neck pain. *Spine*. 2000;25(9):1148-1156. doi:10.1097/00007632-200005010-00017

[16] Burton AK, Waddell G, Bartys S, Main CJ. Screening to identify people at risk of long-term incapacity: A Conceptual and Scientific review.; 2003. https://www.researchgate.net/publication/28578348_Screening_to_identify_people_at_risk_of_long-term_incapacity_a_conceptual_and_scientific_review. Accessed October 21, 2024.

[17] Dionne CE, Von Korff M, Koepsell TD, Deyo RA, Barlow WE, Checkoway H. Formal education and back pain: a review. *Journal of Epidemiology & Community Health*. 2001;55(7):455-468. doi:10.1136/jech.55.7.455

- [18] Waddell G. Occupational health guidelines for the management of low back pain at work: evidence review. *Occupational Medicine*. 2001;51(2):124-135. doi:10.1093/occmed/51.2.124
- [19] Géhin E. Bourdieu Pierre, La distinction, critique sociale du jugement. *Persée*. https://www.persee.fr/doc/rfsoc_0035-2969_1980_num_21_3_5027.
- [20] Boltanski L. Les usages sociaux du corps. *Annales Histoire Sciences Sociales*. 1971;26(1):205-233. doi:10.3406/ahess.1971.422470
- [21] Defrance J. Sociologie du sport - *Éditions La Découverte*. 6th ed. Repères; 2011. <https://shs.cairn.info/sociologie-du-sport--9782707169501?lang=fr>.
- [22] Nicaise S, Mennesson C, Bertrand J. Les inégalités sociales de santé dès la petite enfance : le rôle du sport. *Empan*. 2020;n° 118(2):22-31. doi:10.3917/empa.118.0022
- [23] Mennesson C, Bertrand J, Court M. Forger sa volonté ou s'exprimer: les usages socialement différenciés des pratiques physiques et sportives enfantines. *Sociologie*. 2016;7(4):393-412. <https://shs.cairn.info/revue-sociologie-2016-4-page-393?lang=fr>.
- [24] Cavalli S. Modèle de parcours de vie et individualisation. *Gérontologie et société*. 2006;30(4):55-69. <https://shs.cairn.info/revue-gerontologie-et-societe1-2007-4-page-55?lang=fr>.
- [25] Darmon M. La notion de carrière: un instrument interactionniste d'objectivation. *Politix*. 2008;82(2):149-167. doi:10.3917/pox.082.0149.
- [26] Colinet S. Carrière de la personne atteinte de maladie chronique et éducation thérapeutique. *HAL-Open Science*. <https://hal.science/hal-03288668/document>. Published 2010. Accessed October 21, 2024.
- [27] Mbarga J, Foley RA, Pichonnaz C, Ancey C. Trajectoires de personnes souffrant de lombalgie chronique : ruptures et reconstructions après un programme de rééducation. *Santé Publique*. 2020;Vol. 32(1):19-28. doi:10.3917/spub.201.0019
- [28] Perrin C. Analyse des relations entre le rapport aux APS et les conceptions de la santé. *Staps*. 2015;Hors-série(HS):131-141. doi:10.3917/sta.hs01.0131

- [29] Perrin C, Mino J. Activité physique adaptée et enjeux normatifs dans les trajectoires de maladie chronique: aux points de rencontre des approches interactionnistes et de la philosophie de Georges Canguilhem. *Corps*. 2019;17(1):103-112. doi:10.3917/corp1.017.0103.
- [30] Tousignant-Laflamme Y, Martel MO, Joshi A, Cook C. Rehabilitation management of low back pain – it’s time to pull it all together! *Journal of Pain Research*. 2017;Volume 10:2373-2385. doi:10.2147/jpr.s146485
- [31] Tagliaferri SD, Angelova M, Zhao X, et al. Artificial intelligence to improve back pain outcomes and lessons learnt from clinical classification approaches: three systematic reviews. *Npj Digital Medicine*. 2020;3(1). doi:10.1038/s41746-020-0303-x
- [32] Kendell M, Beales D, O'Sullivan P, Rabey M, Hill J, Smith A. The predictive ability of the STarT Back Tool was limited in people with chronic low back pain: a prospective cohort study. *Journal of Physiotherapy*. 2018;64(2):107-113. doi:10.1016/j.jphys.2018.02.009
- [33] Lam OT, Strenger DM, Chan-Fee M, Pham PT, Preuss RA, Robbins SM. Effectiveness of the McKenzie Method of Mechanical Diagnosis and Therapy for Treating Low Back Pain: Literature Review with Meta-analysis. *Journal of Orthopaedic and Sports Physical Therapy*. 2018;48(6):476-490. doi:10.2519/jospt.2018.7562
- [34] Lamichhane B, Jayasekera D, Jakes R, et al. Multi-modal biomarkers of low back pain: A machine learning approach. *Neuroimage Clin*. 2021;29:102530. doi:10.1016/j.nicl.2020.102530
- [35] Santana AN, de Santana CN, Montoya P. Chronic Pain Diagnosis Using Machine Learning, Questionnaires, and QST: A Sensitivity Experiment. *Diagnostics (Basel)*. 2020;10(11):958. Published 2020 Nov 17. doi:10.3390/diagnostics10110958
- [36] Tack C. Artificial intelligence and machine learning | applications in musculoskeletal physiotherapy. *Musculoskeletal Science and Practice*. 2018;39:164-169. doi:10.1016/j.msksp.2018.11.012

- [37] Robinault L, Niazi IK, Kumari N, Amjad I, Menard V, Haavik H. Non-Specific Low Back Pain: An Inductive Exploratory Analysis through Factor Analysis and Deep Learning for Better Clustering. *Brain Sciences*. 2023;13(6):946. doi:10.3390/brainsci13060946
- [38] Simula AS, Malmivaara A, Booth N, Karppinen J. A classification-based approach to low back pain in primary care – protocol for a benchmarking controlled trial. *BMC Family Practice*. 2020;21(1). doi:10.1186/s12875-020-01135-8
- [39] Jenssen MDK, Bakkevoll PA, Ngo PD, et al. Machine Learning in Chronic Pain Research: A Scoping review. *Applied Sciences*. 2021;11(7):3205. doi:10.3390/app11073205
- [40] Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *The Lancet*. 2011;378(9802):1560-1571. doi:10.1016/s0140-6736(11)60937-9
- [41] Sanchis-Sánchez E, Lluch-Girbés E, Guillart-Castells P, Georgieva S, García-Molina P, Blasco JM. Effectiveness of mechanical diagnosis and therapy in patients with non-specific chronic low back pain: a literature review with meta-analysis. *Brazilian Journal of Physical Therapy*. 2020;25(2):117-134. doi:10.1016/j.bjpt.2020.07.007
- [42] Rabey ML. Multidimensional Patient Profiles in Chronic Non-Specific Axial Low Back Pain - Subgrouping and Prognosis. <https://espace.curtin.edu.au/handle/20.500.11937/46>. Published 2015.
- [43] Jiang N, Luk KDK, Hu Y. A machine learning-based surface electromyography topography evaluation for prognostic prediction of functional restoration rehabilitation in chronic low back pain. *Spine*. 2017;42(21):1635-1642. doi:10.1097/brs.0000000000002159
- [44] Rasmussen-Barr E, Lindqvist C, Östhols S, Boström C. Are patient reported outcome measures (PROMs) useful in low back pain? Experiences of physiotherapists in primary health care in Sweden. *Musculoskeletal Science and Practice*. 2021;55:102414. doi:10.1016/j.msksp.2021.102414
- [45] Gauvin MG, Riddle DL, Rothstein JM. Reliability of clinical measurements of forward bending using the modified Fingertip-to-Floor method. *Physical Therapy*. 1990;70(7):443-447. doi:10.1093/ptj/70.7.443

- [46] Ito T, Shirado O, Suzuki H, Takahashi M, Kaneda K, Strax TE. Lumbar trunk muscle endurance testing: An inexpensive alternative to a machine for evaluation. *Archives of Physical Medicine and Rehabilitation*. 1996;77(1):75-79. doi:10.1016/s0003-9993(96)90224-5
- [47] Biering-Sørensen F. Physical measurements as risk indicators for Low-Back Trouble over a One-Year period. *Spine*. 1984;9(2):106-119. doi:10.1097/00007632-198403000-00002
- [48] Charles M. Épidémiologie de l'inactivité physique et de la sédentarité. Un outils d'évaluation : le questionnaire d'activité physique ONAPS-PAQ. <https://dumas.ccsd.cnrs.fr/dumas-03447001v1>. Published February 18, 2021.
- [49] Lawlis GF, Cuencas R, Selby D, McCOY CE. The development of the Dallas Pain questionnaire. *Spine*. 1989;14(5):511-516. doi:10.1097/00007632-198905000-00007
- [50] Coste J, M LPJ, Berge E, Delecouillerie G, Paolaggi JB. Validation française d'une échelle d'incapacité fonctionnelle pour l'évaluation des lombalgies (EIFEL). https://www.lissa.fr/rep/articles/PF_4846036. Published 1993.
- [51] Lesage FX, Martens-Resende S, Deschamps F, Berjot S. Validation of the General &i>Health Questionnaire<i> (GHQ-12) adapted to a work-related context. *Open Journal of Preventive Medicine*. 2011;01(02):44-48. doi:10.4236/ojpm.2011.12007
- [52] Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*. 1993;52(2):157-168. doi:10.1016/0304-3959(93)90127-b
- [53] van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*. 2011;45(3):1-67. doi:10.18637/jss.v045.i03
- [54] Hotelling H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*. 1933;24(6):417-441. doi:10.1037/h0071325
- [55] Benzécri JP. *L'Analyse des Données. Volume II: L'Analyse des Correspondances*. Dunod; 1973.

- [56] van der Maaten L, Hinton G. Visualizing data using t-SNE. *Journal of Machine Learning Research*. 2008;9:2579-2605. <https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf>
- [57] McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for dimension reduction. *arXiv preprint*. 2018. <https://arxiv.org/abs/1802.03426>
- [58] Ward JH Jr. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*. 1963;58(301):236-244. doi:10.1080/01621459.1963.10500845
- [59] Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise (DBSCAN). *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. 1996:226-231. <https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf>
- [60] Campello RJGB, Moulavi D, Sander J. Density-based clustering based on hierarchical density estimates (HDBSCAN). In: *Advances in Knowledge Discovery and Data Mining (PAKDD)*. Springer; 2013:160-172. doi:10.1007/978-3-642-37456-2_14
- [61] Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1977;39(1):1-22. <https://www.jstor.org/stable/2984875>
- [62] Akaike H. A new look at the statistical model identification (AIC). *IEEE Transactions on Automatic Control*. 1974;19(6):716-723. doi:10.1109/TAC.1974.1100705
- [63] Schwarz G. Estimating the dimension of a model (BIC). *Annals of Statistics*. 1978;6(2):461-464. doi:10.1214/aos/1176344136
- [64] Zou H, Hastie T. Regularization and variable selection via the Elastic Net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2005;67(2):301-320. doi:10.1111/j.1467-9868.2005.00503.x
- [65] Breiman L. Random forests. *Machine Learning*. 2001;45(1):5-32. doi:10.1023/A:1010933404324
- [66] Chen T, Guestrin C. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016:785-794. doi:10.1145/2939672.2939785

[67] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*. 2017;30:4765-4774. <https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf>

Conflit d'intérêts

L'auteur déclare qu'il n'existe aucun conflit d'intérêts lié à cette étude. Cependant, cette recherche est financée par l'Association Nationale de la Recherche et de la Technologie (ANRT), dans le cadre d'une convention CIFRE. Ce financement n'a exercé aucune influence sur la conception, la réalisation ou l'interprétation des résultats de l'étude.