

FICHE SUJET DE THESE

Sujet N° (à remplir par l'ED) :	FINANCEMENT : <input checked="" type="checkbox"/> Demandé <input type="checkbox"/> Acquis	Origine du financement :
Titre de la thèse : Machine-learning based identification of patient-specific risk factors for the early osteoarthritis management		3 mots-clés : Machine-learning, osteoarthritis, predictive modelling
Unité/équipe encadrante : LATIM/IMAGINE		
Directeur de thèse : Pr Romuald SEIZEUR		N° de tél : 06 07 63 53 57 Mail : romuald.seizeur@univ-brest.fr
<p><u>Contexte socioéconomique et scientifique (env. 10 lignes) :</u> Selon l'Organisation mondiale de la santé (OMS), l'arthrose était la seizième cause d'années vécues avec un handicap en 1990, la onzième en 2010, la quatrième en 2020, et elle continuera à augmenter au cours des prochaines années. Avec les effets combinés du vieillissement et de l'obésité croissante, le nombre de personnes touchées par l'arthrose augmentera d'environ 50% au cours des 20 prochaines années. Aujourd'hui, il n'y a aucun remède et les thérapies existantes visent uniquement à réduire les symptômes. Le nombre d'opérations de décompression nerveuse (colonne vertébrale) ou de remplacement articulaire (articulations périphériques) a donc augmenté de manière exponentielle ces dernières années. Ce phénomène ne s'explique pas seulement par les changements épidémiologiques. Les indications sont élargies chez les patients plus jeunes dont les symptômes sont moins graves. Cela entraîne des coûts nettement plus élevés en raison du taux d'échec plus élevé, des niveaux de satisfaction plus faibles et du nombre plus élevé de révisions. Récemment, plusieurs experts ont souligné la nécessité de disposer de stratégies de prévention pertinentes pour ralentir, voire inverser le processus d'arthrose, en particulier, la prévention secondaire qui implique une prise en charge précoce des patients dont la gravité de la maladie est faible et qui sont sujets à une détérioration de leur mode de vie en raison de l'évolution certaine de la maladie.</p>		
<p><u>Hypothèses et questions posées (env. 8 lignes) :</u> Pour avoir un programme de prévention secondaire efficace, le défi principal est d'identifier pour un patient donné les facteurs de risque qui pourraient être atténués avec un programme de prévention approprié. La progression de l'arthrose est très hétérogène, multifactorielle. C'est pourquoi l'anticipation de l'évolution de cette maladie chez un patient donné est un défi important. Les approches basées sur l'apprentissage automatique, et plus particulièrement celles basées sur l'apprentissage profond, ont récemment montré des résultats très prometteurs en matière de prévision, principalement en raison de leur capacité à analyser des ensembles de données importants et multidimensionnels. Mais ces approches ne sont toujours pas utilisées aujourd'hui en raison de plusieurs problèmes non résolus tels que la généralisabilité, c'est-à-dire la capacité à faire des prédictions en utilisant des données provenant de sources autres que celles utilisées pour l'entraînement, et la transparence de la décision qui est liée à la démystification du processus de raisonnement de ces algorithmes. En outre, la plupart des approches de prédiction proposées ont été développées pour estimer la progression de la maladie uniquement, principalement parce qu'elles ont été initialement conçues pour le développement de médicaments uniquement. Avec cette thèse, nous voulons aller plus loin en utilisant des modèles de prédiction pour identifier les facteurs de risque les plus pertinents pour un patient donné atteint d'arthrose.</p>		
<p><u>Grandes étapes de la thèse (env. 12 lignes) :</u> La thèse sera divisée en deux tâches principales : (1) la collecte et la structuration des données, permettant (2) la construction et la validation d'un modèle de prédiction dédié à l'identification des facteurs de risque. La première étape de la thèse sera axée sur les données, avec pour objectif de développer un algorithme de modélisation prédictive généralisable. La principale limite aujourd'hui est que la majorité des solutions de prédiction existantes basées sur l'apprentissage automatique ont été entraînées et validées en utilisant un seul ensemble de données. Pour surmonter ce problème, nous voulons entraîner et valider notre algorithme en utilisant deux grandes cohortes indépendantes publiquement accessibles qui comprennent des données médicales concernant respectivement 4 796 et 3 026 patients. La deuxième étape de la thèse consistera à élaborer, à partir de ces données, un nouveau modèle d'apprentissage profond permettant la détermination objective et transparente de ces facteurs de risque spécifiques aux patients. Comme les algorithmes de prise de décision basés sur l'apprentissage profond sont souvent considérés comme une boîte noire, il est important de rendre le processus de raisonnement de ces algorithmes plus transparents pour les utilisateurs. Mais en raison du comportement multifactoriel de la maladie et des données d'entrée multimodales, les approches existantes sont souvent limitées. Nous voulons développer de nouvelles méthodes de transparence qui pourraient être particulièrement pertinentes pour les données multimodales en mettant conjointement en évidence les caractéristiques de toutes les données des patients. L'interprétation combinée de toutes ces caractéristiques mises en évidence améliorera la transparence et augmentera ainsi la confiance des utilisateurs.</p>		
<p><u>Compétences scientifiques et techniques requises par le candidat (2 lignes) :</u> Le candidat devra posséder de solides compétences théoriques et pratiques en mathématiques appliquées, traitement d'images, apprentissage statistique et profond. Il/elle devra également posséder de solides bases en développement logiciel (C++ ou Python). Une bonne maîtrise de l'anglais est indispensable. Le/la candidat(e) devra également montrer un fort intérêt pour le domaine de la santé.</p>		
<p><u>3 publications de l'équipe d'accueil relatives au domaine (5 dernières années) :</u> [1] Dardenne G, Pluchon JP, Letissier H, Guezou-Philippe A, Gerard R, Lefèvre C, Stindel E. Accuracy and precision of an Ultrasound-based device to measure the pelvic tilt in several positions. Journal of Ultrasound in Medicine 2019. doi:10.1002/jum.15141. [2] Dardenne G, Dib Z, Poirier N, Letissier H, Lefèvre C, Stindel E. What is the best hip center location method to compute HKA angle in computer-assisted orthopedic surgery? In silico and in vitro comparison of four methods. Orthop Traumatol Surg Res 2019;105(1):55-61.doi: 10.1016/j.otsr.2018.11.011. [3] Villa M, Dardenne G, Nasan M, Letissier H, Hamitouche C, Stindel E. FCN-based approach for the automatic segmentation of bone surfaces in ultrasound images. Int J Comput Assist Radiol Surg 2018;13(11):1707-1716.doi:10.1007/s11548-018-1856-x.</p>		
<p><u>Collaborations nationales et internationales :</u> Ce sujet s'inscrit dans le cadre de nouveaux projets qui ont pour but de développer de nouvelles solutions technologiques permettant la prise en</p>		

charge globale et précoce de patients atteint d'arthrose. Au niveau local nous travaillons sur ces aspects en étroite collaboration avec le service de chirurgie orthopédique, de Rhumatologie et de Neurochirurgie du CHRU de Brest. Sur le plan national, le LaTIM est également un des membres fondateurs du LabEx National CAMI (Computer Assisted Medical Interventions), et participe activement à l'infrastructure Nationale FLI (France Life Imaging) avec le pilotage d'actions relatives à l'imagerie interventionnelle et multimodales. Sur le plan régional, le laboratoire est porteur de la plateforme Bretonne PIMATGI, volet Therafonc et coordonne conjointement la FHU Technologies pour la santé.