

**Nevoret C, Laurendeau C, Tessier C, Voinot C, Zins M, Goldberg M**

**Apports et limites du "machine learning" dans la prédiction du changement du stade de sévérité de l'asthme en France : une analyse du Système national des données de santé (SNDS)**

**VIIIe Congrès National Conjoint ADELFF-EMOIS, 31 mars - 1 avril 2022, Dijon**

#### **RÉSUMÉ**

INTRODUCTION - Prédire l'évolution de la maladie permettrait d'améliorer la prise en charge des patients et de réduire le fardeau de la maladie. La base de données médico-administratives du SNDS pourrait permettre de construire des modèles prédictifs de l'évolution de la sévérité des maladies et d'orienter des politiques de prise en charge des patients. OBJECTIF - Déterminer dans quelle mesure le SNDS pourrait identifier un modèle prédictif de l'aggravation de l'asthme. METHODES - L'étude repose sur une exploitation des données de la cohorte CONSTANCES chaînées au SNDS. L'ensemble des patients asthmatiques en 2017 ont été inclus. Le stade GINA de sévérité de la maladie a été estimé chaque mois pour chaque patient par un algorithme reposant sur des consommations médicamenteuses. L'aggravation du stade GINA en 2017 a été analysée à travers différents modèles prédictifs issus du « machine learning » (régression logistique, RF, SVM, KNN, réseau de neurones) dont les performances ont été comparées via la sensibilité, la spécificité, l'« accuracy » et la matrice de confusion. Les variables prédictives ont été identifiées par des cliniciens au sein des consommations de soin ou du questionnaire CONSTANCES en 2016. RESULTATS - Au total, 5007 patients asthmatiques âgés en moyenne de 47 ans ont été inclus dans l'étude (hommes : 44,7 %). Le stage GINA était de : non traités : 37 %, stade 1 : 22 %, stade 2 : 3 %, stade 3 : 16 %, stade 4 : 20 % et stade 5 : 2 %. Au cours de l'année, 34,4 % des sujets ont eu une aggravation du stade. Les modèles de ML permettent d'obtenir une bonne prédiction de cette évolution dans 70 % des cas (« accuracy ») entre 0,65 et 0,75 quel que soit le modèle testé. Des différences notables sont observées en termes de sensibilité, entre 0,45 (KNN) et 0,82 (SVM) et de spécificité entre 0,58 (SVM) et 0,86 (RF). DISCUSSION / CONCLUSION - Tous les modèles testés sont comparables en termes de performances globale et très variables sur la sensibilité et la spécificité du modèle.

**MOTS CLÉS : -**

**RÉSUMÉ PUBLIÉ DANS :** [Revue d'Epidémiologie et de Santé Publique. 2022; 70\(1s\):S6. doi: 10.1016/j.respe.2022.01.064.](https://doi.org/10.1016/j.respe.2022.01.064)

**INFORMATIONS COMPLÉMENTAIRES, [ICI](#)**