

Introduction :

- L'accès aux bases de données (BDD) de santé (SNDS) permet des recherches sur des thématiques particuliers. L'utilisation des statistiques conventionnels pour ces BDD est possible mais elle donne une vision limitée des possibilités explorations de ces bases.
- Nous avons réalisé une étude sur l'agressivité en fin de vie chez les patients atteints de cancer du poumon et décédés à l'hôpital à partir de la base PMSI. L'agressivité en fin de vie a été définie par la présence d'un des 4 critères suivant :
 - Réalisation d'une chimiothérapie dans les 14 jours précédant le décès
 - 1 hospitalisation pour symptômes dans les 30 derniers jours
 - Hospitalisation en USI et/ou en réanimation
 - Passage en Soins Palliatifs < 3 jours précédant dans le décès.
- Nous avons inclus dans notre cohorte 79 749 patients décédés en 2010/2011 ou en 2015/2016. Au moins 57% des patients de cette cohorte présentés au moins 1 facteur d'agressivité.
- Les statistiques classiques tel que l'analyse univariée et multivariée permet de rechercher des déterminants associés à des soins potentiellement « agressives ». En analyse multivariée, on retrouve que l'âge, le sexe, l'ancienneté du cancer, la présence de certains comorbidités, le type de structure hospitalière, l'index de déprivation sociale et la densité de la commune de résidence sont des facteurs indépendants après ajustement sur l'agressivité des soins de la fin de vie.

• Néanmoins, ce type d'analyse limite le nombre de variable à inclure dans l'analyse multivariée pour permettre de tirer des conclusions. En présence d'un nombre important de variable, les nouvelles techniques dites « intelligence artificielle » peuvent permettre d'apporter des informations complémentaires comme identifiées des profils-types de patients qui n'exprime pas d'indicateur d'agressivité ou au contraire surexprimé des critères d'agressivité.

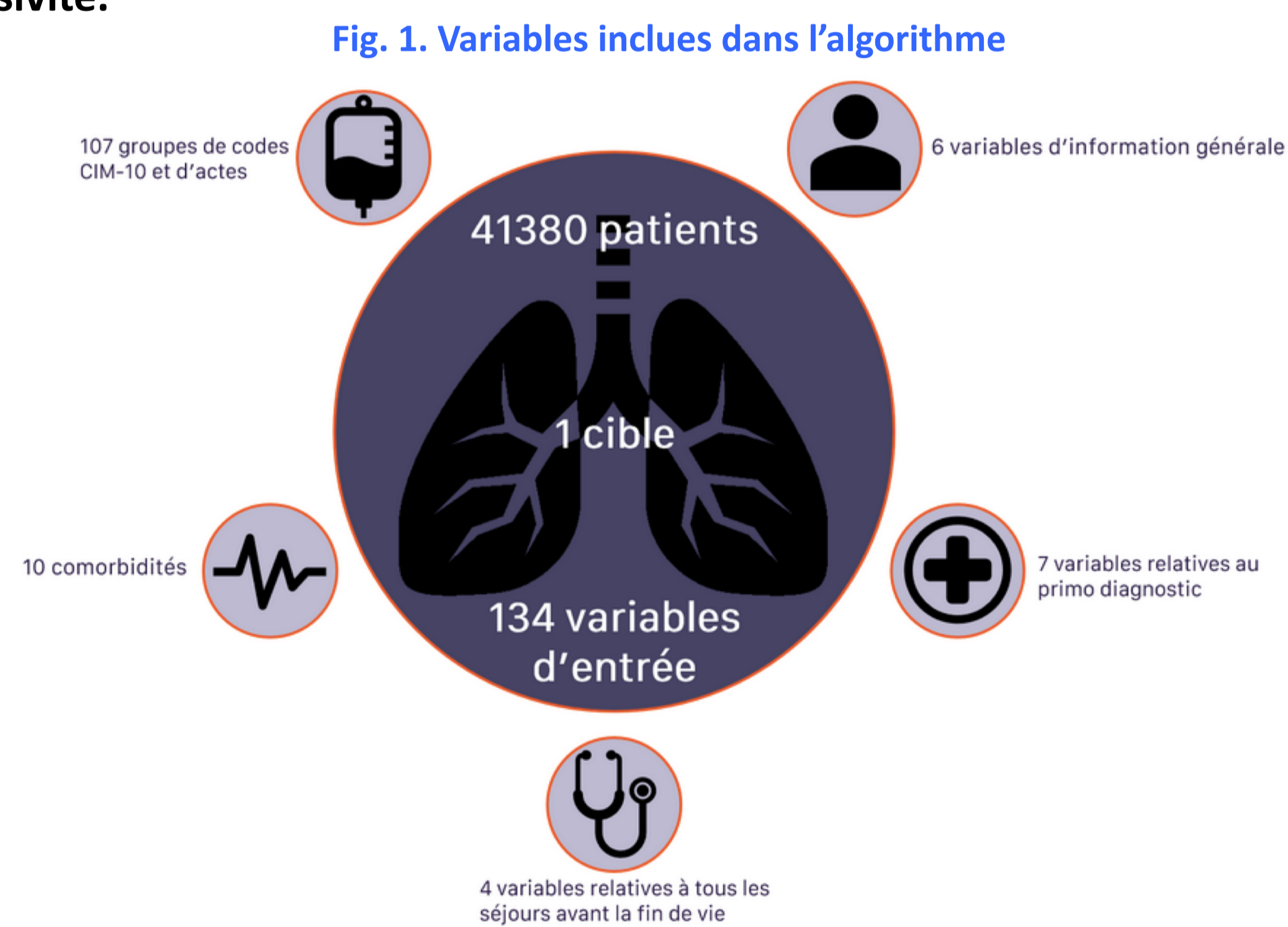


Fig. 1. Variables incluses dans l'algorithme

Méthodes

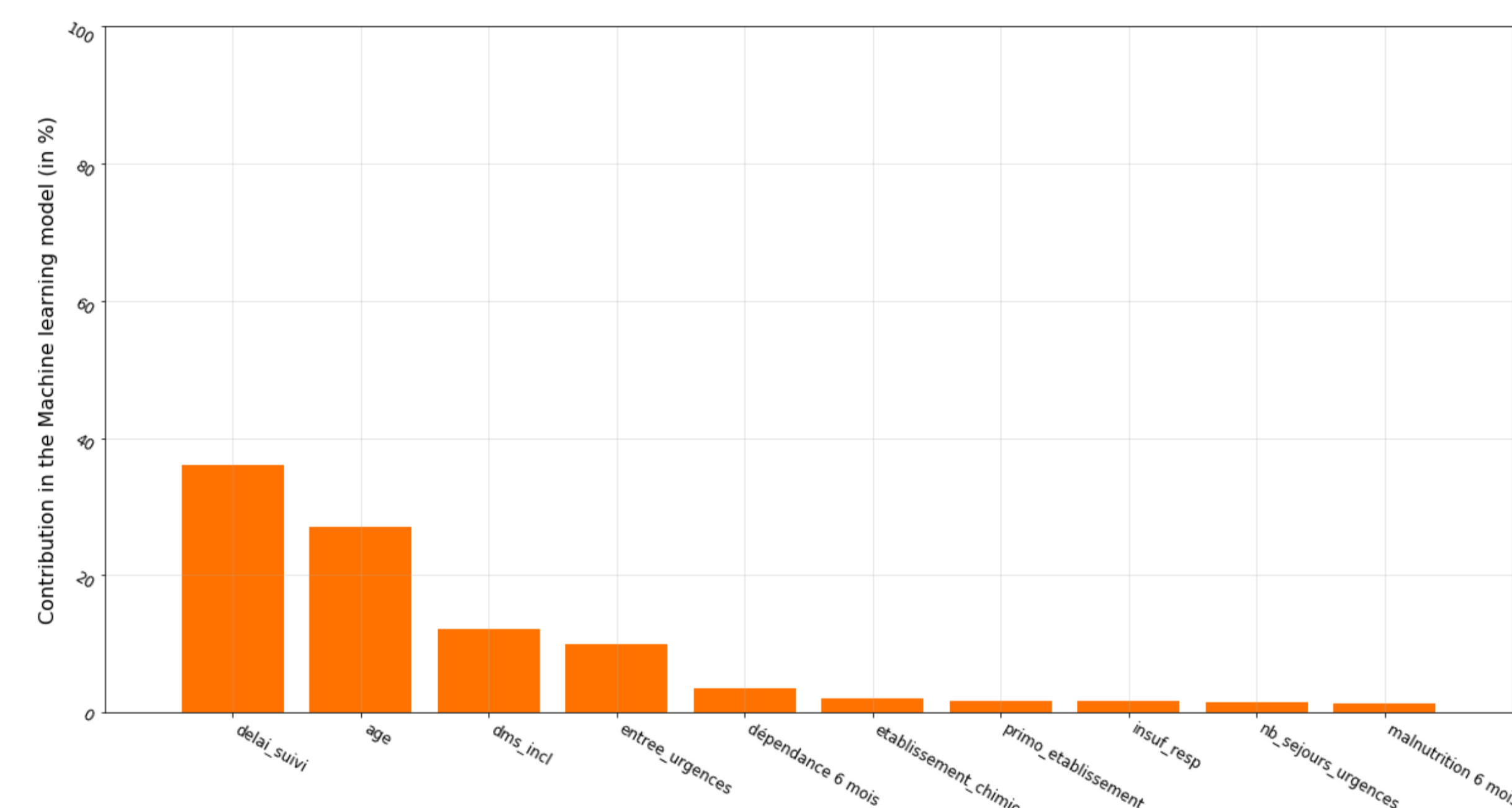
Représentation graphique

- Le graphe « Sunburst »: Il est constitué d'une succession de cercles concentriques. En partant du centre, chaque étape de l'algorithme de décision (= une opération de séparation) est représentée par un niveau divisé en deux parties selon les caractéristiques des patients.
- Les « Feature Importance »: C'est un résultat clé associé aux arbres de décisions. Il s'agit d'un diagramme en barres où chacune d'entre elles représente une variable d'entrée. La hauteur des barres est proportionnelle à l'influence de la variable dans l'algorithme de décision.

Critères de qualité

- La pureté d'un profil de patients représente le pourcentage de la classe majoritaire dans ce profil. Plus la pureté est élevée, plus le résultat d'un profil est significatif. Par exemple, si un profil comporte 100 patients dont 60 ont un score égal à 1 et 40 un score égal à 0, alors la pureté sera de 60% et la classe retenue pour le profil sera « score égal à 1 ».
- Pour ceux issus d'une régression, l'indicateur qui remplace la pureté est l'erreur absolue moyenne (EAM). Si cette erreur est trop grande, le niveau du Sunburst est grisé.

Fig. 2. Feature importance pour l'agressivité de fin de vie pour la classification binaire



Résultats

Classification binaire

- On peut identifier des profils-types de patients avec un score d'agressivité à 0 ou des patients avec un score d'agressivité ≥1.
- Les variables les plus discriminantes sont : délai diagnostic-décès (35%), âge au diagnostic (25%), durée du 1ère séjour (15%), primo-diagnostic via les urgences (12%). Ils permettent de définir des Profil-Patient les plus prédictifs d'avoir ou non des soins agressifs en fin de vie. (Fig. 2.)
- Le TOP 3 des profil-types est décrit dans la Fig. 3.

Top 3 des profils à retenir pour la classification

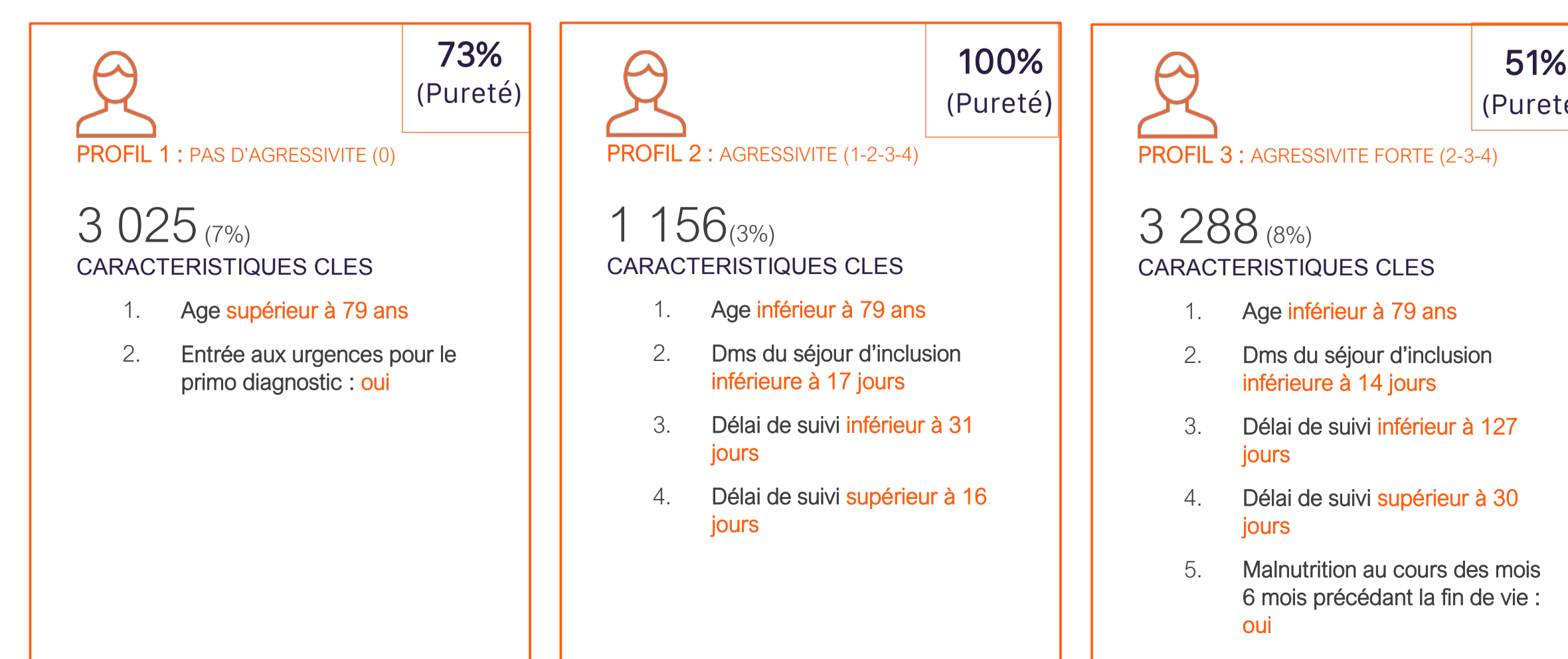
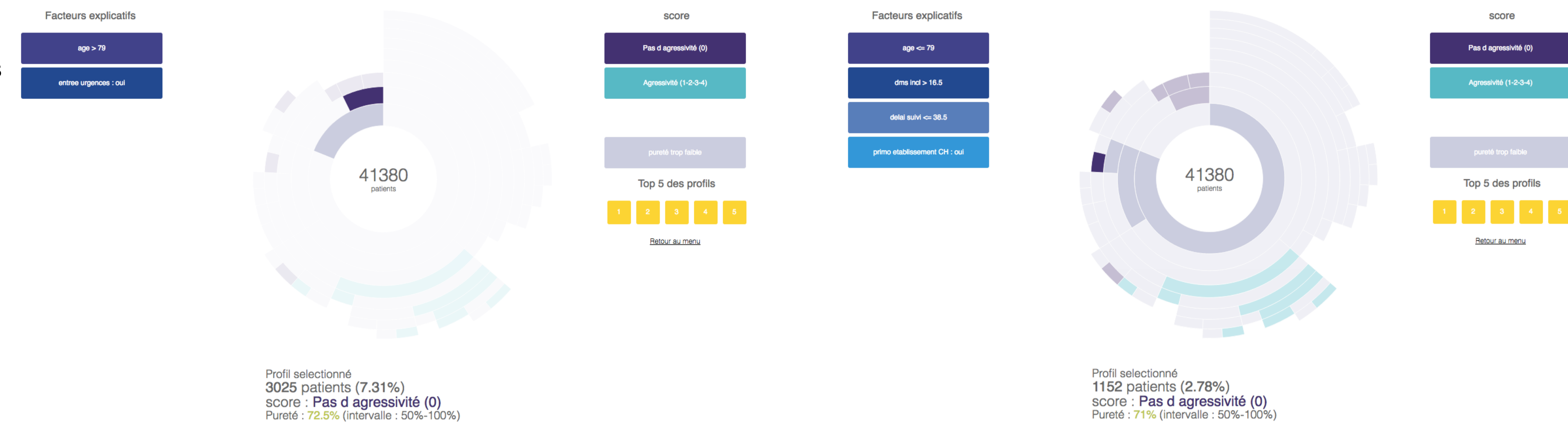


Fig. 4. Profil-type de Patient qui n'exprime pas de facteurs d'agressivités



Profil-type

- Profil-type de patients qui n'expriment pas d'indicateurs d'agressivité avec une bonne pureté (>50%) (Fig 4.)
- Profil-type de patients qui ont une forte probabilité d'avoir des critères d'agressivités (Fig 5.)

Fig. 5. Profil-type de Patient qui surexprime de facteurs d'agressivités



Discussion et Conclusions

- La technique de Marching Learning est une méthode complémentaire des statistiques conventionnelle car elle permet d'établir des profils-types de patients non réalisables avec l'analyse multivariée. On peut essayer de focaliser nos actions de prévention sur certains profils de patients.
- Dans cette analyse de l'agressivité de la fin de vie chez les patients atteints de cancer du poumon, certains résultats confortent l'analyse uni et multivariée :
 - Age : plus un patient est âgé, moins il a de chance de subir un traitement agressif au cours de sa fin de vie.
 - Délai de suivi : plus ce temps de suivi est important, plus l'agressivité est forte
- On arrive également à faire ressortir d'autres variables d'intérêts grâce à ces techniques.
- Plus la durée du séjour d'inclusion est importante, plus le risque d'agressivité sera faible. Ce résultat est d'autant plus vrai si le patient est entré par les urgences. Cela démontre que lorsque le séjour permettant de repérer le cancer du poumon subi des complications, il est plus facile d'appréhender la fin de vie d'un patient et donc d'agir en conséquence pour ne pas le surcharger en traitements superflus.
- L'utilisation des techniques « intelligence artificielle » sur les BDD ouvre des nouvelles perspectives de recherche.
- Pour notre étude, les résultats montrent des résultats intéressants mais encore limité par le nombre de variable disponible dans le PMSI. Le chaînage des différentes bases de données du SNDS (SNIRRAM, EGB) permettrait d'alimenter le nombre de variable à analyser.
- L'application de cette technique sur des cohortes prospectives de cancer du poumon (base ESME-poumon) pour permettre d'améliorer les résultats avec l'utilisation de variable cliniques ou biologiques non disponible dans le SNDS.