

Prédiction de la mortalité à court terme après la pose d'un défibrillateur cardiaque implantable

Utilisation de réseaux de neurones auto-encodeurs sur les données du SNDS

Introduction

Les données de remboursement non cliniques du SNDS (Système National des Données de Santé) sont prometteuses, mais présentent des challenges pour les études pharmaco-épidémiologiques. En particulier, l'étude des **parcours des patients**, c'est-à-dire la séquence temporelle de soins (hospitalisations, médicaments, consultations, etc.), présente un intérêt majeur pour déterminer des **facteurs prédictifs**.

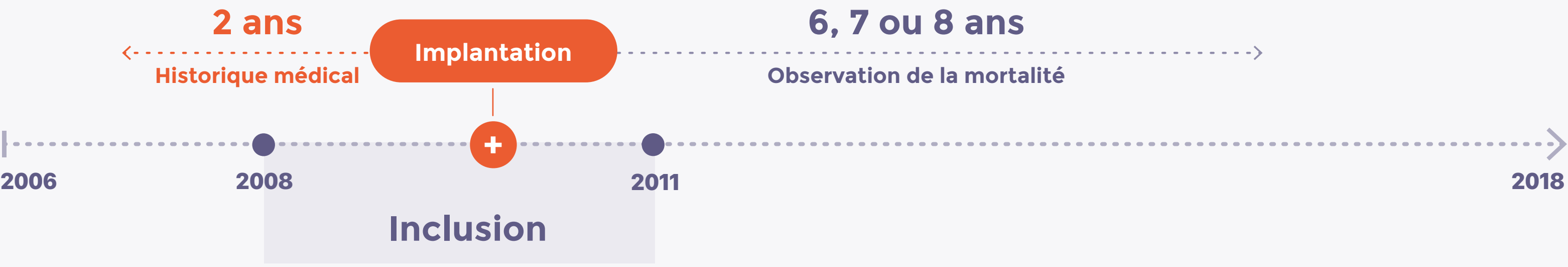
Cependant, les séquences temporelles sont hautement combinatoires, le parcours de chaque patient étant unique. **Des algorithmes prédictifs performants et explicables sont donc nécessaires.**

Ce travail se concentre sur la **prédiction d'un risque lié au parcours de soin** et sur l'identification de ses **facteurs prédictifs**.

Méthodologie

Design de l'étude

Les patients sont inclus sur l'implantation d'un DAI (Défibrillateur Automatique Implantable) entre 2008 et 2011. Les deux années précédant l'inclusion sont extraites pour obtenir l'historique de soin des patients. Une période d'observation de la mortalité a lieu suite à l'inclusion : la **mortalité à court terme** se définit grâce à une fenêtre temporelle définie par la durée de vie estimée du DAI (respectivement 6, 7 ou 8 ans selon que le DAI soit simple, double ou triple chambre).



Modélisation

Une méthodologie analytique utilisant des **réseaux de neurones auto-encodeurs** est utilisée pour obtenir des prédictions explicables à partir de données brutes.

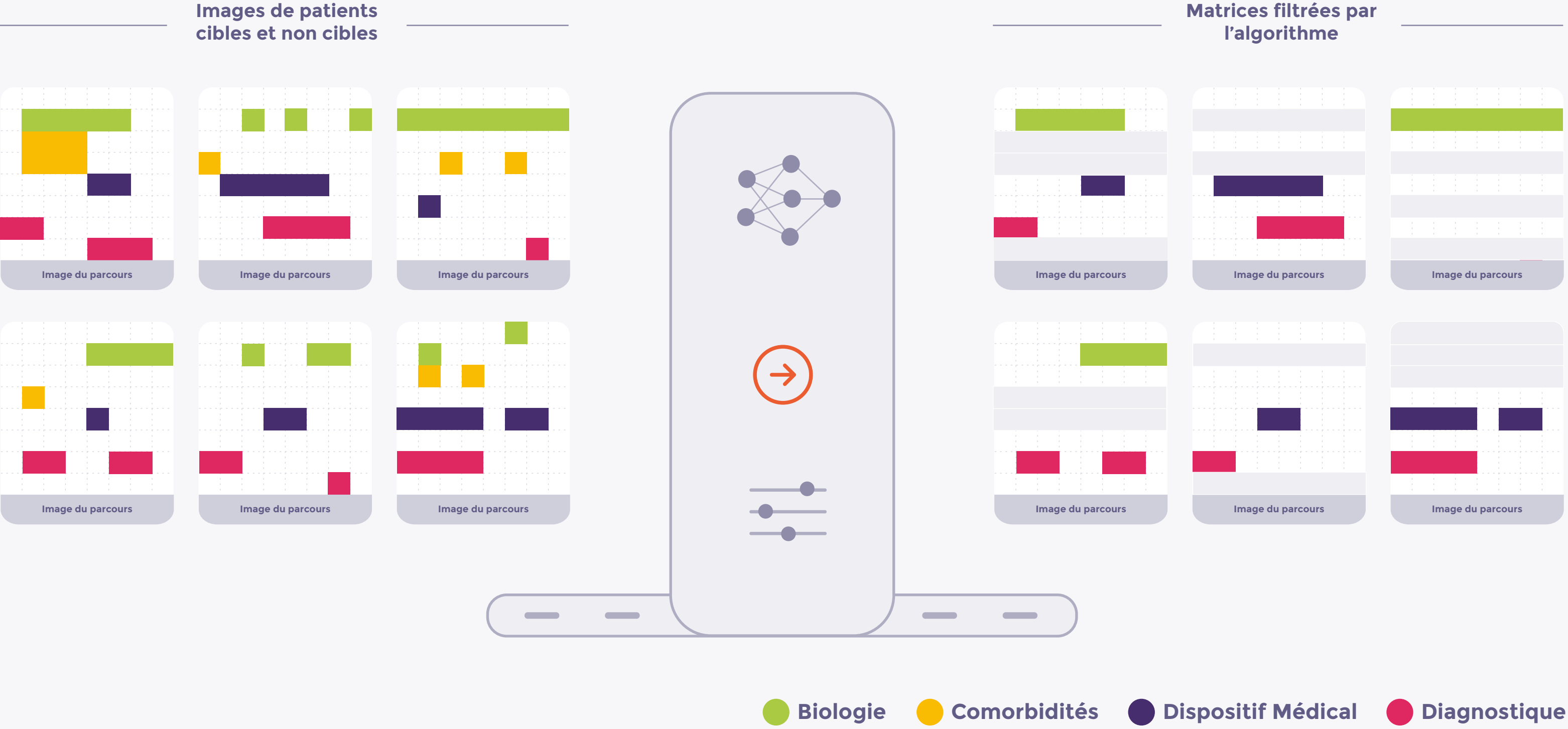
L'étude de cas porte sur la prédiction de la mortalité à court terme après la pose d'un défibrillateur cardiaque implantable avec les données pré et post-implantation des patients.

Nous modélisons les parcours de soin bruts des patients en générant des **images** de ces parcours qui soient utilisables dans l'algorithme et expliquant les **facteurs de risques**.



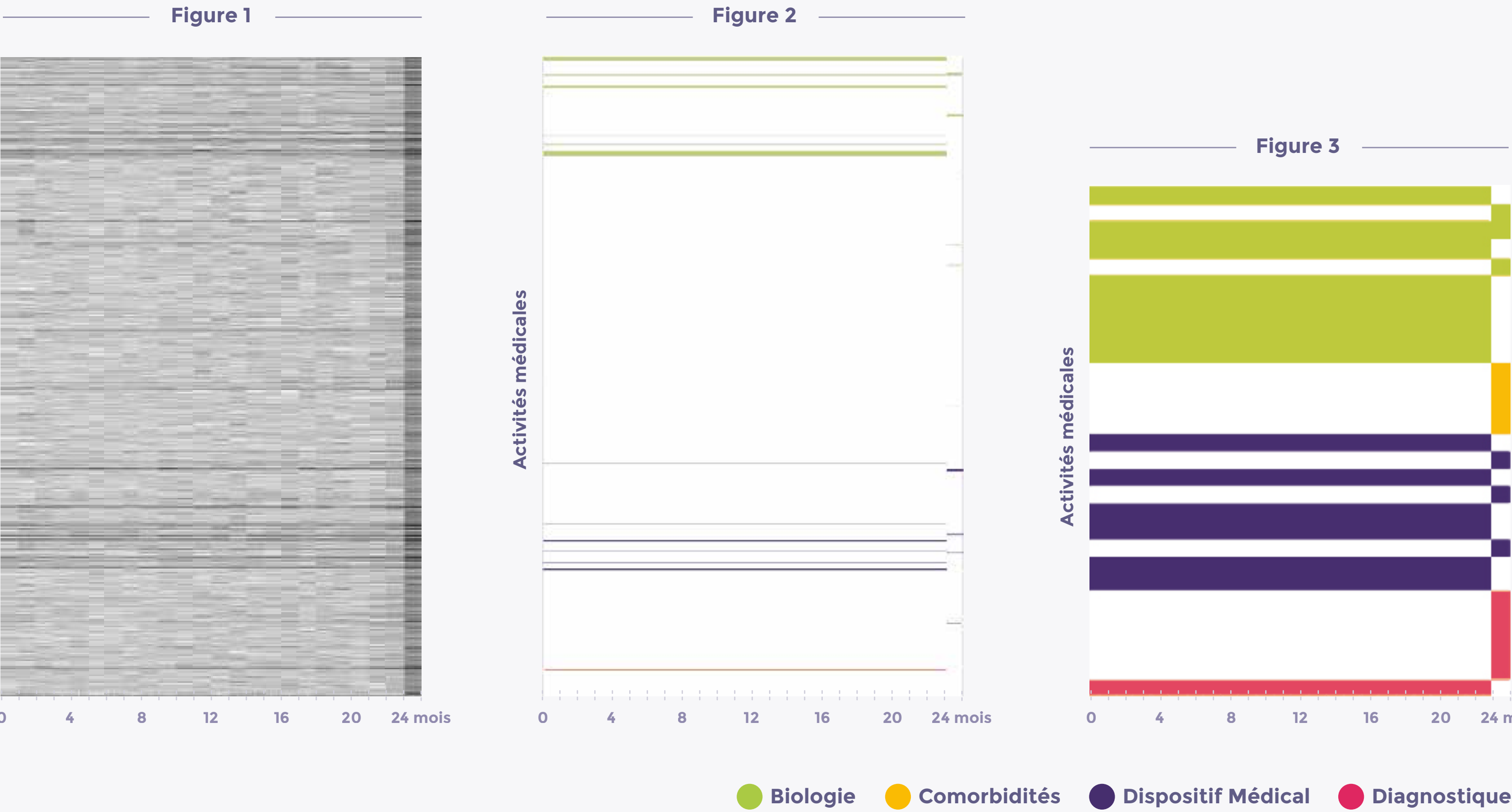
Algorithme

L'algorithme est entraîné à ne reconstruire que les événements typiques des patients cibles, il met donc en avant **les événements décisifs** pour la prédiction de mortalité à court terme.



Résumé des événements prédicteurs trouvés

Nous générons l'**agrégation des événements marquants** de tous les parcours (Fig.1). Puis, nous filtrons et ne mettons en avant **seulement les événements les plus importants** (Fig 2 et Fig 3).



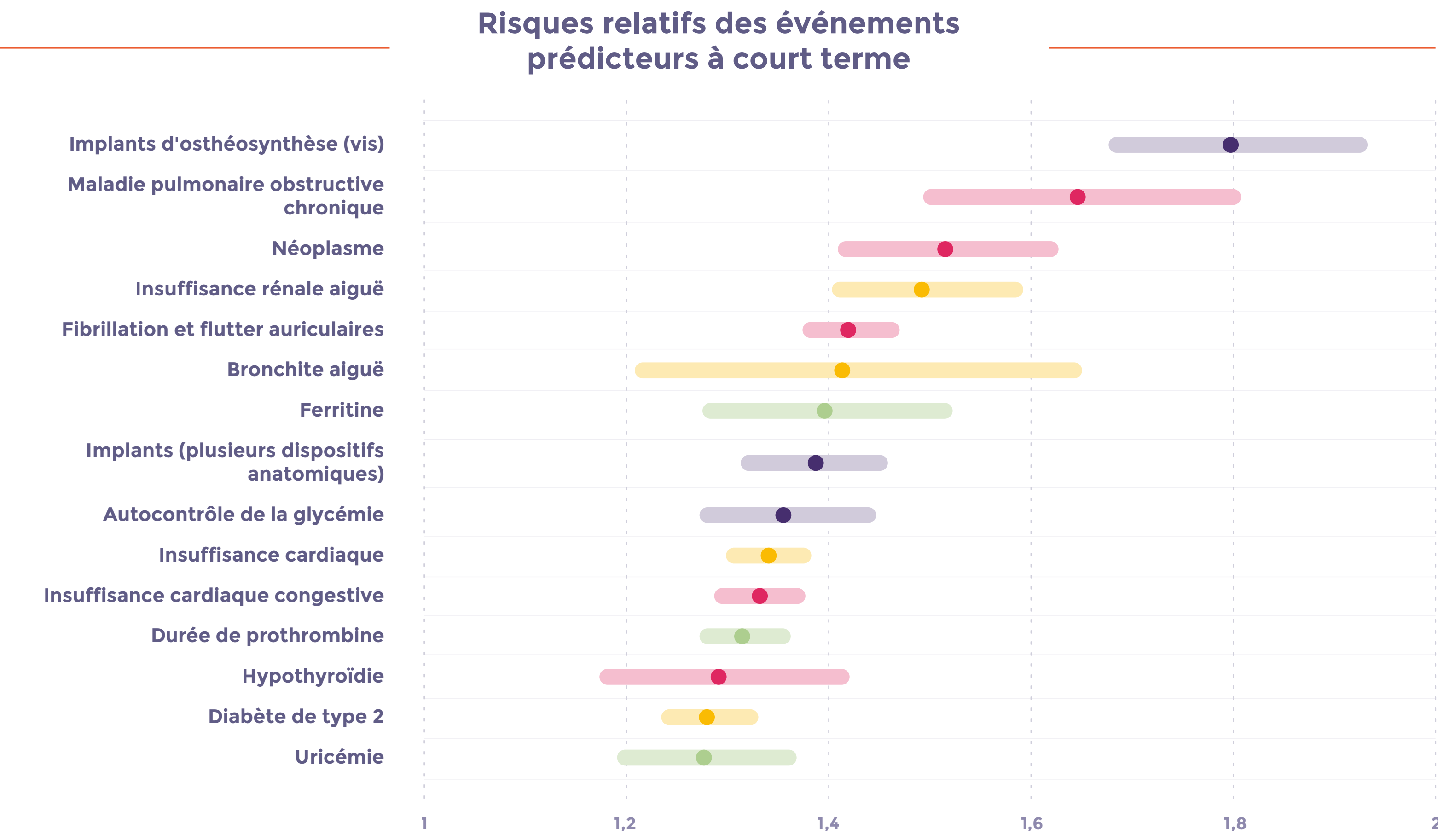
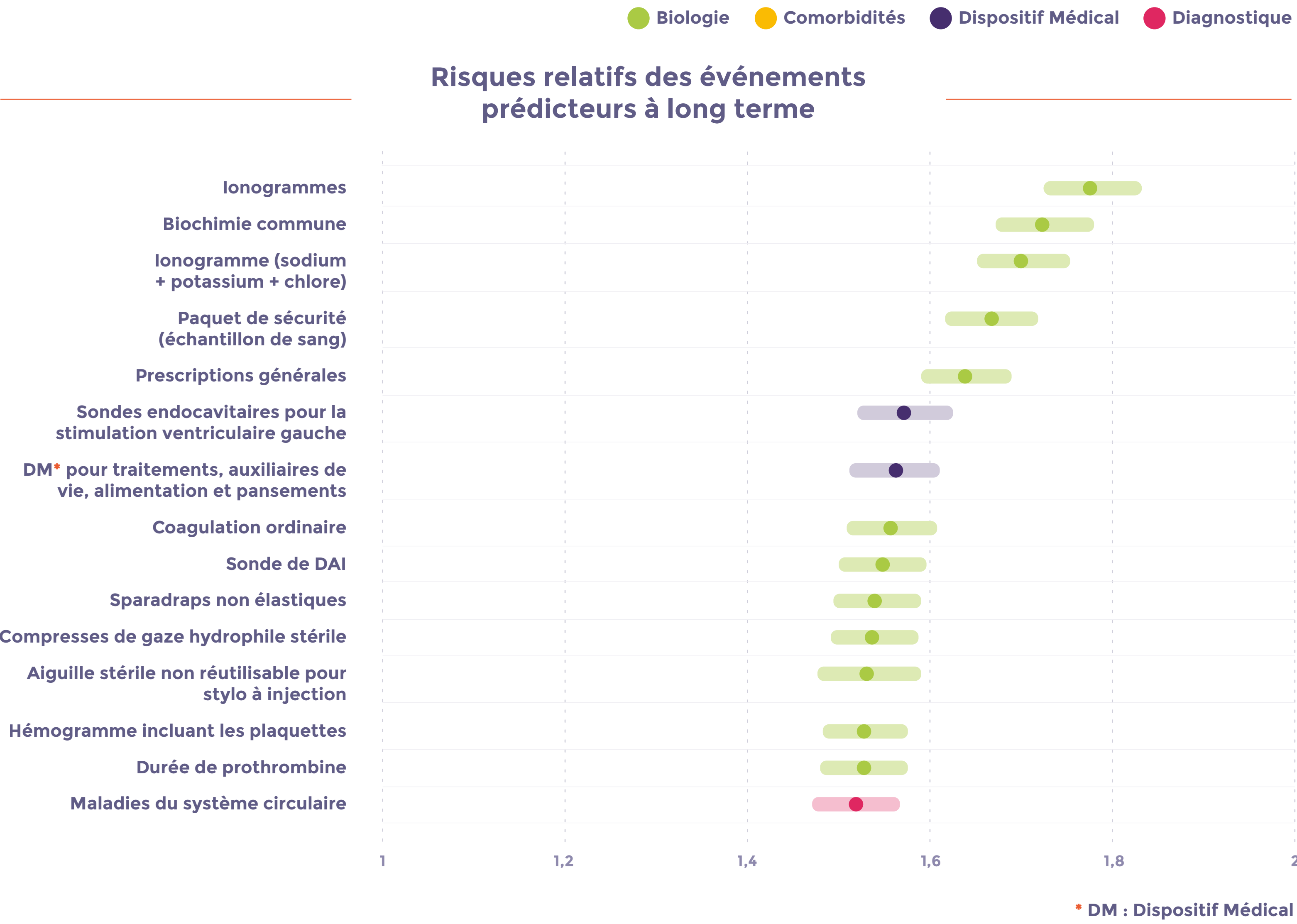
Résultats

Facteurs de risque

15 facteurs de risque ont été jugés significatifs, avec des risques relatifs compris entre 1,2 et 1,8.

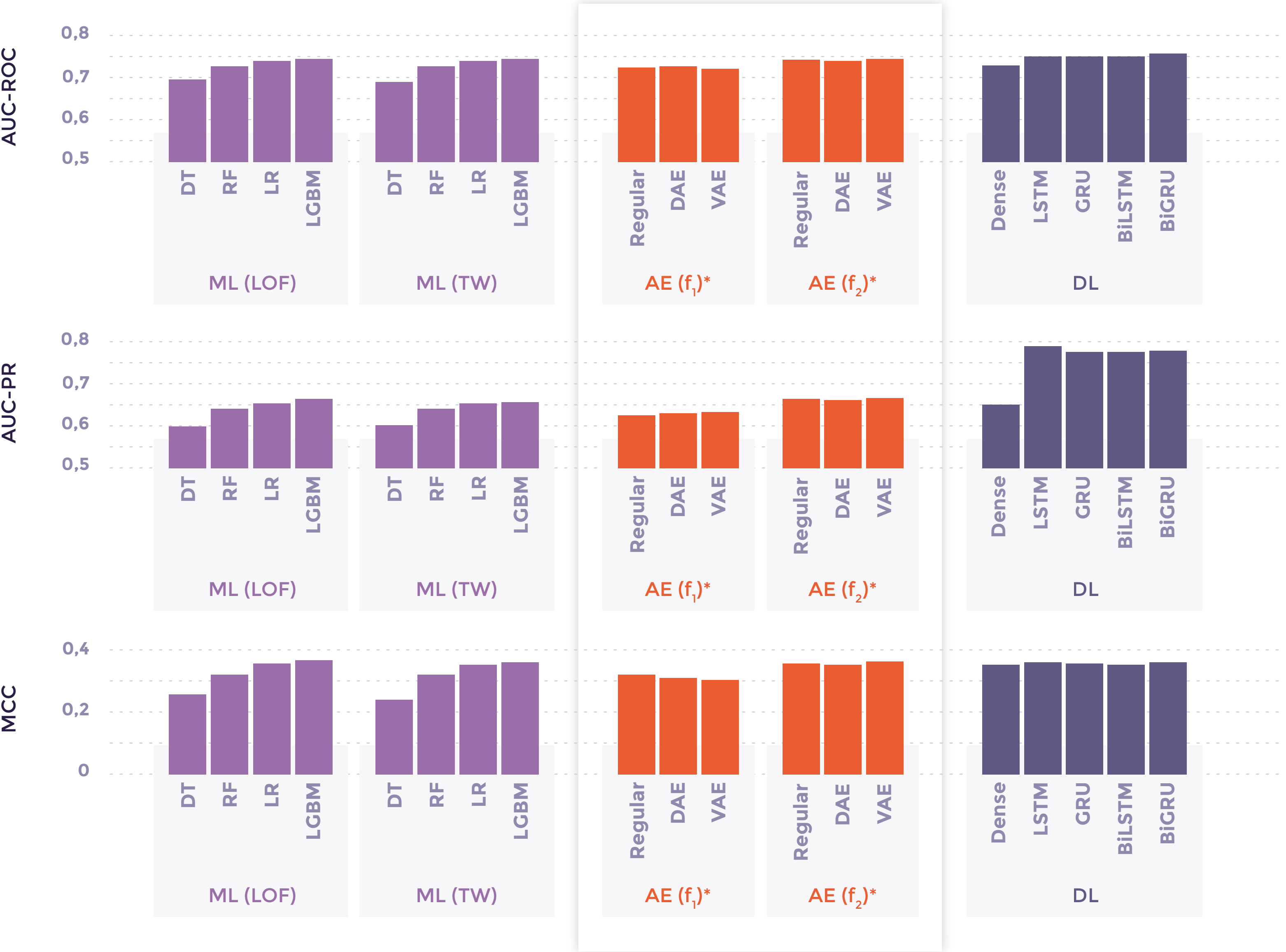
Les examens biologiques récurrents, les dispositifs médicaux associés et la fréquence des hospitalisations arrivent en tête. Les **facteurs de risque identifiés**, confrontés à une **expertise médicale**, permettent d'identifier les **patients à risque**. Deux types d'événements à risque sont identifiés :

- Long terme : Ils peuvent avoir lieu n'importe quand dans le temps
- Court terme : Ils ont lieu dans la fenêtre temporelle précédant l'implantation



Performances

Les performances globales de prédiction **sont comparables aux modèles** de machine learning « boîte noire », tout en étant explicables.



Conclusion

Cette modélisation des données avec auto-encodeurs est générique et peut être employée sur d'autres cas d'étude pour identifier les facteurs prédictifs à partir du SNDS dès que la prédiction d'un risque est liée à un historique d'événement de soins et que l'explicabilité est requise. Cette contribution est utile pour les décideurs qui ont besoin d'analyses prédictives étayées par une extraction claire des connaissances des données.