

## Retour d'expérience de Nicolas Glatt, président fondateur de CLINITYX, membre AFCROs, sur le cas d'une étude SNDS appariée en chirurgie endo-vasculaire :

**« Notre projet d'étude SNDS appariée a nécessité un nombre important de compétences. »**

### **Quel a été le point de départ du projet ?**

A la demande de la HAS, la Société Française de Cardiologie (SFC) a suivi toutes les implantations de valves aortiques sous cutanées en France depuis la première implantation en 2010. Au moment où l'on commence à travailler avec la SFC, cette cohorte France TAVI regroupe 45 000 patients porteurs d'une valve aortique, implantée par 50 centres, avec une tendance à 12 000 inclusions par an. Elle nous demande de digitaliser la collecte des données qui sont entièrement collectées par eCRF pour faire un suivi individuel orienté sur les événements indésirables et la consommation de soins. Nous avons découvert le SNDS à cette occasion pour proposer un nouveau paradigme à la SFC.



### **Quelles ont été les options choisies ?**

La cardiologie interventionnelle, qui implante les valves, fait appel à deux éditeurs très présents sur le territoire, si bien que les données collectées sont assez homogènes dans les systèmes hospitaliers. Le SNDS devait pouvoir nous permettre de compléter le registre, avec les données de suivi et de complications, sans solliciter les investigateurs. Avec la CNAM, nous avons choisi la méthode d'appariement probabiliste renforcée par un acte traçant assez fort et un dispositif médical onéreux. L'objectif nous a été donné d'apparier 90% des patients. Pour traiter les données, nous avons monté notre infrastructure pour héberger le SNDS dans une « bulle sécurisée » conforme au référentiel de sécurité du SNDS. On s'est ensuite lancé dans la procédure d'accès aux données.

### **Comment se passe la phase d'extraction des données ?**

France TAVI a été le premier projet pour lequel on a demandé une extraction à la CNAM donc il a été regardé de façon attentive par la CNIL. Un délai de 13 mois a été nécessaire pour récupérer l'extraction. Le CESREES a émis un avis favorable rapidement mais côté CNIL, nous avons dû produire une étude de risque poussée et un gros travail d'architecture technique pour démontrer que nous étions sur un même niveau de sécurité que la CNAM, gestionnaire du SNDS « natif ». En pratique, une première zone récupère les données qui viennent des systèmes hospitaliers ou des centres investigateurs et, une fois par an, nous transmettons le fichier d'appariement à la CNAM. Globalement, nous avons accès pour chaque patient à 12 ans de données, deux ans avant la date d'implantation et jusqu'à 10 ans après. La CNAM nous transmet le SNDS Fils qui correspond aux patients qui sont inclus. Avec du recul aujourd'hui, on peut dire que l'appariement probabiliste n'a pas été évident au départ et a nécessité 5 itérations pour réussir un premier appariement validé par la SFC. Lorsque l'on a récupéré un SNDS fils, nous avons eu une longue phase de data engineering pour retraiter les incohérences



et médicaliser le SNDS. Une fois ce travail réalisé, on s'aperçoit que le nombre de patients « perdus de vue » a significativement diminué, ce qui est exceptionnel pour une cohorte de cette taille, et on a ainsi retrouvé une mortalité cohérente. L'apport du SNDS dans le registre se révèle donc au final positif.

### **Comment envisagez-vous la suite du projet ?**

Aujourd'hui, c'est un projet qui a vocation à vivre puisqu'il a une autorisation de la CNIL pendant 7 ans, au cours desquels il y a aura 5 renouvellements d'appariement. L'usage de ces données est double : certaines sont destinées au Ministère de la Santé et à la HAS et d'autres à des publications scientifiques. Le comité scientifique de la SFC va sélectionner un certain nombre d'études que nous allons mettre en place. Avec l'enjeu de les industrialiser. Deux publications ont déjà été faites sur France TAVI : l'une sur la méthode d'appariement et l'autre sur les anticoagulants et le TAVI. En 2021, 10 études ont été retenues par le comité scientifique ancillaire pour continuer d'exploiter cette BDD. Cela vise des projets d'étude variés sur des analyses descriptives sur l'impact de tel ou tel critère sur le devenir du patient ou des projets basés sur le *machine learning* avec des modèles prédictifs pour prédire l'AVC ou la mortalité en fonction de critères patient.

### **Peut-on dire que la démarche est un succès ?**

Pour la SFC, c'est un vrai succès sur un plan scientifique et au niveau de la communication. Il y a la volonté d'élargir à presque tous les registres. Côté transparence et information, nous avons obtenu l'autorisation d'apparier sans informer nominativement les patients inclus, mais nous devons créer un site d'information collectif. Appréhendé ainsi, le projet a nécessité un nombre de compétences importantes : médicale (comité scientifique), réglementaire, infrastructure (une personne et demi à plein temps chez nous), des épidémiologiques, *data engineers*, *data scientists* ou statisticiens. Le SNDS est formidable pour son exhaustivité, très peu médicalisé, avec des biais de codage considérables... Par conséquent, il revêt encore une part de secret que l'on n'a pas réussi à totalement élucider. En outre, l'approche des entrepôts de données est intéressante et mériterait d'être soutenue par l'ensemble des acteurs de l'écosystème.