



HAL
open science

Confier une décision vitale à une machine

Clément Henin

► **To cite this version:**

Clément Henin. Confier une décision vitale à une machine. Réseaux : communication, technologie, société, 2021, L'action publique au prisme de la gouvernamentalité numérique, 225, pp.187 - 213. 10.3917/res.225.0187 . hal-03551778

HAL Id: hal-03551778

<https://hal.science/hal-03551778>

Submitted on 1 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CONFIER UNE DÉCISION VITALE À UNE MACHINE

Clément Henin

Résumé : depuis 2018, un algorithme développé par l'agence de la biomédecine établit l'ordre de proposition des greffons cardiaques, ressource vitale en quantité insuffisante pour les malades qui en ont besoin. Alors qu'ils sont amputés d'une partie de leur autonomie professionnelle, les médecins acceptent globalement ce système centralisé et automatique, bien que l'attribution semble imparfaite, car il est jugé préférable à une attribution humaine qui était largement manipulée par les centres de greffe. Le terrain permet d'observer de nouvelles pratiques dans les centres de greffe qui s'approprient le fonctionnement de l'algorithme afin de déployer des tactiques pour minimiser l'impact de l'algorithme et de participer aux discussions sur les évolutions du système automatique.

Mots-clés : algorithme de décision, gouvernementalité algorithmique, choix tragique, attribution d'organes

Les candidats à la greffe de cœur souffrent d'une insuffisance cardiaque mettant en jeu leur pronostic vital et ont épuisé toutes les alternatives thérapeutiques.¹ Compte tenu de la forte mortalité de ces patients et de leur qualité de vie dégradée, la greffe est considérée comme leur meilleur traitement. L'opération chirurgicale de transplantation pose en général peu de problèmes et la difficulté réside surtout dans l'obtention d'un greffon, car cela implique qu'une personne compatible et consentante se trouve en état de mort encéphalique. Au cours de l'année 2016, l'Agence de la biomédecine (ABM) a comptabilisé deux fois plus de candidats inscrits sur liste d'attente que de greffons disponibles (ABM, 2016), il s'agit donc d'une situation de « choix tragique » (Calabresi et Bobbitt, 1984) dans laquelle un traitement est disponible, mais pas en quantité suffisante pour répondre aux besoins de la population. Afin d'obtenir au juste moment un greffon compatible, les médecins transplantateurs des centres de greffe (côté receveur) ont intérêt à collaborer en mettant en place une liste nationale². Ce milieu reste pourtant compétitif puisque les médecins, désireux de soulager les souffrances de leur patient, souhaitent s'approprier la ressource, quitte à employer des stratégies illicites au détriment des autres patients en attente (Thibaut, 2013).

Dans ce contexte, le système assurant l'attribution de cette ressource précieuse joue un rôle central dans la prise en charge. Depuis 2018, c'est le « score cœur », un algorithme développé par une équipe de l'ABM, qui établit l'ordre de proposition des greffons, succédant à un système dans lequel les médecins transplantateurs avaient une grande autonomie dans le choix des patients à greffer. L'attribution d'un traitement vital rare est un problème qui dépasse la décision purement médicale (Childress, 2001). Persad *et al.* (2009) proposent au moins trois catégories de principes éthiques pour aborder spécifiquement le problème de l'attribution des

¹ Je tiens à remercier toutes les personnes qui ont participé à ce travail et l'ont rendu possible. En premier lieu je remercie les médecins et personnels soignants qui m'ont accordé un peu de leur temps précieux ainsi que les membres de l'équipe des scores d'attribution de l'agence de biomédecine. Je remercie également Julien Barrier et Julie Henry qui m'ont aiguillé dans la phase d'entretien et m'ont apporté des conseils précieux pour la rédaction. La version finale de cet article doit aussi beaucoup à Elsa Zotian, Nathanael Colin-Jaeger et les membres du comité de rédaction que je remercie aussi. Enfin, je remercie Sophie Quennelle pour m'avoir donné l'idée de cette étude originale dans le milieu de la greffe cardiaque.

² Certains font même fait le choix de regroupements internationaux, comme les membres du réseau Eurotransplant qui regroupe huit pays européens.

ressources médicales : équité, prioritarisme et utilitarisme³. Les auteurs précisent qu'aucun de ces principes n'est suffisant. Il n'est pas acceptable d'attribuer la ressource à celui qui souffre le plus s'il est de toute façon condamné, pas plus qu'il n'est acceptable de l'attribuer à celui qui en profitera le plus si ce dernier peut attendre. Seule une combinaison de ces principes peut satisfaire nos exigences morales. Pour calculer le score qui établit la priorité d'accès à un greffon, l'algorithme applique un ensemble de règles et de formules aux caractéristiques numérisées du donneur et du receveur. Ce traitement informatique combine des modèles prédictifs issus des données historiques de la greffe et des règles éditées « à la main » pour obtenir un compromis entre les critères éthiques listés plus haut.

Les sciences sociales ont déjà montré que les algorithmes, loin d'être des moyens techniques neutres, configurent les possibles (Rouvroy et Berns, 2013). D'autres auteurs ont par ailleurs décrit la question plus générale du pouvoir associé à la quantification et aux statistiques (Desrosières, 2010 ; Supiot, 2015). Afin de replacer la question algorithmique dans un contexte historique plus large, nous nous appuyons sur les travaux de Theodore Porter (1996) qui a étudié comment au XXe siècle certaines professions ont été forcées par des pressions extérieures à abandonner la discrétion et l'autonomie que leur conféraient leurs professions au profit de standards, de règles ou d'objectifs chiffrés. Par ailleurs, l'analyse de la normalisation des pratiques en cancérologie proposée par Castel et Merle en 2002 permet également de replacer notre analyse. Alors que ces normes semblent aller à l'encontre des médecins en réduisant notamment leur autonomie via des pratiques d'objectivation et de rationalisation du jugement médical, les sociologues montrent que les médecins les acceptent et participent même à l'initiation de la démarche. La possibilité d'adapter ces normes, de se les approprier et d'apprendre à les maîtriser pour les employer à leur avantage semble justifier cette acceptation. Les méthodologies propres aux sciences sociales ont régulièrement été employées afin d'étudier comment certains dispositifs algorithmiques configurent les réseaux d'acteurs dans les comportements en ligne (Badouard *et al.*, 2016), la police dite « prédictive »

³ L'attribution équitable donne à chacun les mêmes chances d'accéder à la ressource. Le tirage au sort ou la règle « premier arrivé, premier servi » en sont des exemples. La priorité aux plus mal-lotés ou prioritarisme donne la priorité au patient qui souffre le plus ou au patient le plus jeune qui a le moins profité de ses années de vie en bonne santé. Enfin, le calcul utilitariste tente de maximiser les retombées positives en sauvant le plus grand nombre de vies ou le plus grand nombre d'années de vie ; les patients qui ont le meilleur pronostic sont privilégiés.

(Benbouzid, 2018 ; Brayne et Christin, 2020), la justice et le journalisme (Christin, 2017) ou le moteur de recherche de Google (Cardon, 2013). Cependant, alors même que les promesses liées aux algorithmes dans le domaine médical reçoivent une attention particulière de la part des pouvoirs publics (Villani, 2018), ces derniers trouvent difficilement leur place à l'hôpital (Elwyn *et al.*, 2013), pour des raisons plus organisationnelles et sociologiques que techniques (Yang *et al.*, 2016).

La question de l'allocation des greffons, comprenant l'utilisation de scores et d'algorithmes, a été récemment étudiée par Philippe Steiner (Steiner, 2017 ; Steiner, à paraître). L'auteur aborde le problème de l'attribution des greffons rénaux sous l'angle de l'appariement entre un donneur (offre) et un receveur (demande) avec un fort ancrage dans la littérature économique en faisant notamment référence aux récents travaux de Alvin Roth (2004), récipiendaire en 2012 du prix dit Nobel d'économie pour ces travaux sur l'appariement. Dans ces articles, l'auteur isole la place particulière que l'appariement algorithmique basé sur des données historiques occupe dans le mécanisme d'allocation. D'après l'auteur, le dispositif décrit comme « paternaliste », car laissant peu de place à l'avis des patients, a une forte dimension politique et doit donc pouvoir être contesté par le public. Dans notre article, dont l'objet est l'attribution des greffons cardiaques, nous mettons l'accent sur l'impact de l'introduction d'un outil algorithmique dans le milieu des professionnels de la greffe en nous appuyant sur un travail de terrain, offrant ainsi une approche complémentaire aux travaux de Philippe Steiner.

L'analyse proposée dans cet article s'appuie sur une série de vingt-six entretiens conduits dans quatre centres hospitalo-universitaires français, dont un centre parisien, auprès de médecins transplantateurs et de personnels soignants. Ces entretiens, menés entre juin 2019 et décembre 2019, avaient pour objectif d'aborder la question de la compréhension de l'algorithme ainsi que d'étudier son utilisation. Par ailleurs, des discussions avec l'équipe de conception du score à l'ABM, des observations de réunions à l'ABM et dans les services de greffes ainsi qu'une étude bibliographique ont permis d'enrichir l'analyse.

Je présente d'abord certains éléments qui ont caractérisé l'émergence du score cœur, système algorithmique de décision vitale embarquant de fortes dimensions éthiques. À quel problème le score répond-il ? Comment y

répond-il ? C'est ensuite la réception du score par les centres de greffe qui est étudiée. Je montre que les médecins sont amputés d'un pouvoir de décision et se voient imposer des pratiques avec lesquels ils ne sont pas toujours en accord. Malgré cela le score est bien accepté, non pas pour ces propriétés techniques, mais parce qu'il représente une amélioration par rapport à la situation à laquelle il succède et parce qu'il laisse une certaine marge de manœuvre aux médecins. En effet, des tactiques visant à minimiser l'impact du système d'attribution sont employées. Elles mobilisent la connaissance que les médecins ont du comportement de l'algorithme.

L'ÉMERGENCE DU SCORE CŒUR

Suite aux nombreux dysfonctionnements du système de répartition des greffons antérieur géré par une association de médecins, l'ABM est créée en 2005 pour tenter d'uniformiser les pratiques. Plus tard, le score cœur est déployé pour rationaliser et centraliser l'attribution grâce à un algorithme utilisant les données de soin.

La création de l'agence de la biomédecine

Après la première greffe réussie sur des jumeaux homozygotes en 1954, c'est la découverte de la ciclosporine, traitement immunosuppresseur prévenant le rejet du greffon, qui rend possible l'utilisation de la greffe comme technique de soin dès les années 1980. Rapidement, l'association France-Transplant, créée et administrée par des médecins, met en place une politique d'échanges entre les hôpitaux, puis une liste d'attente nationale en appliquant des règles visant à une distribution équitable des organes (Steiner, 2010). Malgré ces règles, des dysfonctionnements apparaissent tels que la domination des centres parisiens, la présence de patients étrangers sur les listes d'attente ou l'inscription de certains patients dans plusieurs hôpitaux (Déchamp-Le Roux, 1997). Dans un contexte où l'affaire du sang contaminé a déjà entaché la confiance du grand public envers le milieu médical et où l'association n'a pas le pouvoir de contraindre les centres de greffe, l'association France-Transplant est forcée de demander l'intervention de l'État afin d'éviter un nouveau scandale naissant (Coquide, 1993). En 1993, un rapport de l'inspection générale des affaires sociales prend acte de ces

dysfonctionnements ce qui conduit à la création de l'agence de la biomédecine, qui succède à l'établissement français des greffes en 2005.

La régulation de l'activité de transplantation ainsi que l'attribution des greffons sont explicitement inscrites dans les objectifs de cet établissement public qui dispose, à la différence de l'association, de la légitimité administrative pour imposer des règles contraignantes aux centres de greffe. L'organisme mène alors une politique de centralisation et de contrôle des pratiques des centres de greffes. Cela passe par l'instauration de règles strictes applicables à tous les centres et par la mise en place d'un système obligatoire de collecte des données de la greffe dénommé « Cristal ». Comme nous allons le voir, la création de l'ABM et l'instauration de ces mesures ne sont pas sans incidence sur la prise en charge des malades.

L'attribution des greffons cardiaques

Avant l'instauration du score cœur en 2018, l'attribution des greffons cardiaques est principalement faite en deux temps : l'organe est d'abord alloué à l'équipe de greffe la plus proche du lieu de prélèvement et cette équipe l'attribue ensuite à l'un de ses patients selon des critères qu'elle décide. Les médecins de l'équipe ont toute autonomie et discrétion dans ce dernier choix. Entre 2004 et 2018, un régime d'exception, appelé « super urgence », est mis en place pour permettre à certains candidats de profiter d'une priorité nationale d'accès⁴. Cette priorité surpasse l'allocation locale des greffons et donne temporairement à celui qui en bénéficie la première place nationale dans l'ordre de proposition des greffons. Pour qu'il soit considéré en super urgence, le candidat à la greffe doit recevoir de la dobutamine –un médicament utilisé en cas de graves dysfonctions cardiaques– ou son cœur doit être assisté par une circulation extra-corporelle de courte durée (pompe externe remplaçant le travail du cœur lorsque ce dernier n'assure plus sa fonction)⁵. Alors que le deuxième traitement nécessite une opération chirurgicale et s'accompagne d'importants risques de complications, une faible dose de dobutamine peut être administrée à un patient en attente de greffe avec un risque de complication limité. Alors que

⁴ Limité à 48 h renouvelable une fois.

⁵ L'obtention du statut de « super urgence » est théoriquement conditionnée à l'avis d'un membre d'un collège expert, mais en pratique la quasi-totalité des demandes sont acceptées (Dorent *et al.*, 2014).

la dobutamine est un traitement médical visant à améliorer l'état de santé d'un malade, son usage peut donc être détourné par les médecins dans le but d'accéder à la priorité nationale. Il est communément admis que ce système de priorité basé sur la dobutamine a été détourné par les médecins (ou au moins par certains centres) afin de donner la priorité à leurs patients au détriment du reste de la liste. Les statistiques globales montrent d'ailleurs une augmentation régulière du nombre de patients ayant obtenu une priorité nationale entre 2012 et 2017 (ABM, 2017), jusqu'à atteindre une situation où plus de la moitié des greffés a bénéficié du niveau maximum de priorité.

Dans ce système, les médecins des centres de greffe jouissent donc d'une grande autonomie tant dans l'attribution locale que dans l'obtention d'une priorité nationale. L'efficacité et l'équité d'un tel dispositif sont largement remises en question. Les transplantateurs, reconnaissant que le système a permis de diminuer la mortalité sur la liste d'attente, notent en revanche que la mortalité après la greffe a augmenté (Dorent *et al.*, 2014). D'autres soutiennent que le système est inéquitable, car les patients bénéficiant d'une priorité nationale finissent par avoir un risque de mortalité inférieur à ceux qui n'en bénéficient pas, alors qu'il est censé s'agir de patients plus graves (Cantrelle, 2018, p. 87). L'équipe de l'ABM qui se charge de la création du score cœur note que le système ne permet pas d'attribuer efficacement les priorités, qu'il incite les équipes à utiliser excessivement certains traitements qui sont associés à des risques de complication et que le système est mal protégé contre le « gaming » (Jasseron *et al.*, 2017). Certains médecins, dont une partie se retrouve associée à l'équipe de conception, appellent dès 2011 à la mise en place d'un système d'attribution indépendant de la pratique médicale qui jugerait « objectivement » la priorité en se basant sur un score calculé automatiquement (Dorent *et al.*, 2011, *ibid*). Dès le début, l'ambition du score est donc de normaliser l'attribution des greffons afin d'éviter le jeu des acteurs locaux. C'est la mise en place d'un système visant à assurer l'équitable répartition du traitement, en empêchant notamment les jeux locaux des médecins sur le score, qui justifie la rationalisation et la centralisation du processus de décision.

Conception du score cœur

En 2017, l'équipe en charge des scores d'allocation à l'ABM publie un « *Candidate Risk Score* » capable de prédire avec une précision jugée

suffisante par les auteurs le risque de mortalité des patients inscrits sur liste d'attente de transplantation cardiaque tout en limitant la prise en compte des pratiques médicales que les médecins pourraient utiliser pour manipuler le système (Jasseron *et al.*, *ibid*). La construction du modèle s'appuie sur les données historiques des candidats en France entre 2010 et 2013 qui ont été collectées par l'intermédiaire du système d'information Cristal. Les auteurs de l'article identifient les quatre variables les plus souvent associées au décès du patient dans ces données historiques qui sont combinées dans un modèle régressif de Cox (assimilable grossièrement à une régression linéaire ou logistique)⁶. Les variables sont traitées indépendamment et associées à un coefficient. Plus le coefficient est important, plus une modification de la variable affectera la valeur du score de mortalité. Notons que les variables qui peuvent dépendre de la pratique médicale, comme l'usage de dobutamine, sont exclues afin d'obtenir un score *objectif* (Jasseron *et al.*, *ibid*, p. 3). La précision du score est mesurée en comparant les prédictions du modèle à la réalité observée sur un ensemble de données historiques préalablement cachées pour ne pas fausser les résultats.

La publication de ce score, qui est le premier du genre dans le monde, suscite des réactions de la part de la communauté internationale de greffe (Egan, 2017 ; Potena et Khush, 2017). Le score est présenté comme une avancée, bien qu'il semble ne remplir que la moitié de l'objectif (« *Halfway there* »). Selon ces commentateurs, un score complet devrait aussi prendre en compte le risque de décès post-greffe qui seul permet une évaluation utilitaire de la décision d'attribution. En effet, pour minimiser le nombre de décès, il faut à la fois limiter les décès chez les candidats et chez les personnes greffées. En 2018, un score prédictif du risque de décès post-greffe est publié par l'équipe de conception du score (Jasseron *et al.*, 2018), mais la précision de ce score est cette fois jugée insuffisante par les auteurs pour envisager son implémentation complète dans le score d'allocation. Il est pris en compte en demi-mesure pour filtrer les candidats dont le risque prédit de décès un an après la greffe dépasse 50 %, alors que le taux moyen de survie s'établit plutôt autour de 85 %. Cette analyse bibliographique met en lumière un aspect important du score cœur. À cause de l'impossibilité technique d'estimer correctement le risque post-greffe, l'équipe de

⁶ Le décès pouvant survenir à différents instants, il s'agit d'un problème de survie qui impose des modèles plus complexes qu'une régression linéaire standard. Cependant les variables X_1 , X_2 , X_3 et X_4 apparaissent dans la formulation mathématique sous la forme d'une somme avec des coefficients β_1 , β_2 , β_3 et β_4 , comme dans une régression linéaire : $X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + X_4\beta_4$.

conception est forcée de donner une place prépondérante au risque de mortalité sur la liste d'attente. Comme nous le verrons plus loin, la grande priorité accordée aux graves est une critique fréquente dans les centres de greffe.

Sur la base de ce « *Candidate Risk Score* », l'équipe de l'ABM construit le système de répartition des greffons cardiaques : le « score cœur » tel qu'il est déployé en janvier 2018. S'ajoutent aux modèles prédictifs déjà mentionnés un certain nombre de règles issues des connaissances générales du milieu de la greffe, comme la compatibilité en fonction de la morphologie, ou de résultats récents prenant par exemple en considération l'accès réduit à la greffe de certaines sous-populations de candidats. Le score final peut être décrit comme un « patchwork » de modèles prédictifs et de règles ajoutées à la main pour résoudre des problèmes spécifiques. Le score final de l'ABM prend la forme d'un indicateur compris entre 0 et 1102 « points » calculé pour l'ensemble des candidats inscrits sur la liste d'attente de greffe dès qu'un greffon est disponible (Pôle qualité des données, 2019). Le greffon est proposé par ordre décroissant de cet indicateur et l'équipe de greffe a toujours la possibilité de refuser une proposition. Un système d'exception, appelé « composante expert », autorise les centres de greffe à demander une dérogation pour certaines pathologies spécifiques qui sont mal prises en compte par le « *Candidate Risk Score* ». La demande prend la forme d'un dossier évalué par un comité d'experts et peut conduire, s'il est accepté, à l'obtention d'une prime temporaire de points qui accorde une priorité importante à celui qui en bénéficie. Contrairement à Google qui souhaite limiter au maximum les modifications « à la main » de son algorithme Pagerank (Cardon, *ibid*), l'ABM ne réprime aucunement les interventions humaines pour aboutir à la décision finale.

Quelques propriétés sociotechniques du score cœur

Il s'agit d'un système prédictif, centré sur la donnée. Comme le note (Steiner, à paraître), l'appariement des greffons implique une utilisation de la donnée à deux niveaux : comme indicateur des caractéristiques du greffon et du receveur dont l'appariement est envisagé et comme instrument d'évaluation des échanges passés et futurs, puisque les données historiques collectées par Cristal sont utilisées pour évaluer les greffes passées et les potentielles greffes futures. Concrètement, des règles sont imposées sur

l'enregistrement des données. Les médecins doivent fournir ces données suivant une régularité imposée par l'ABM : tous les trois mois en général et tous les trois jours pour les patients urgents. La véracité des données est contrôlée régulièrement dans les centres de greffe. En utilisant des données mesurées, l'ABM affirme une volonté d'un dispositif de décision indépendant du jugement des professionnels.

L'algorithme du score cœur est protéiforme et malléable. Certaines parties du code constituent de véritables modèles prédictifs issus de l'optimisation de paramètres à partir d'une base de données, ce que nous pourrions qualifier d'une méthode d'apprentissage, alors que d'autres sont des règles ajoutées à la main sans appui scientifique. Par exemple, le score de risque ne permettant pas une prédiction suffisamment précise du devenir des mineurs inscrits sur liste d'attente, l'ABM instaure une priorité quasi absolue pour ces patients sans autre justification que son acceptation par la communauté des transplantateurs⁷. Le score est aussi malléable, car, dès sa conception, il a été prévu que l'algorithme pourrait évoluer en fonction des retours des centres de greffe et s'adapter à des cas exceptionnels à la demande des médecins⁸. Les deux mécanismes principaux sont détaillés plus loin.

Un parallèle peut déjà être dressé entre le cas du score cœur et les phénomènes plus anciens décrits par Porter (Porter, 1996). Alors que les comptables du début du XXe siècle jouissaient d'une grande autonomie dans la description des finances des compagnies, des forces extérieures (notamment légitimées par la crise économique de 1930) ont poussé cette profession à se séparer de leur discrétion au profit de standards et de règles objectives fixes. La standardisation imposée aux professionnels, que l'on souhaite mettre en regard de la mise en place de l'algorithme de décision, intervient non pas comme la solution à un problème financier, mais à un problème de confiance. La comptabilité doit devenir un outil de mesure fiable et ne pas reposer sur un corps d'experts en qui il faut avoir confiance. De la même manière, le dispositif de l'ABM opérant une « objectivisation » par des mesures biologiques instaure un système central de contrôle indépendant des localités et des jugements experts. Pourtant, comme nous

⁷ La priorité accordée aux mineurs est une différence notable avec le système Étasunien.

⁸ Cette propriété est d'ailleurs volontiers mise en avant par les concepteurs : « *Le Score Cœur a été conçu pour pouvoir être ajusté en fonction des situations rencontrées ainsi que de l'évolution du profil et de la prise en charge des patients en attente.* » (Richard Dorent dans (ABM, 2019, p. 41)).

l'avons vu, le score cœur autorise un compromis avec la profession par des mécanismes d'adaptation et d'exception. Le dispositif final étant à la fois indépendant et négociable, comme nous le verrons plus en détail à la fin de la deuxième partie.

Enfin, le score est à la fois transparent et opaque. Une politique stricte de transparence a été appliquée par l'ABM puisque j'ai pu recoder, à partir de la documentation disponible en ligne, une version du score dont j'ai pu vérifier avec la collaboration de l'ABM qu'elle obtenait systématiquement les mêmes valeurs que le score réel. Pourtant cette transparence n'est pas synonyme de facilité de compréhension (Burrell, 2016). En effet, le savoir médical encodé dans l'algorithme est structuré dans un ensemble de formules et de règles qui sont parfaitement illisibles, tant pour un informaticien que pour un médecin (Pôle qualité des données, 2019, p.25-27). Cependant, puisque l'algorithme prend la forme de ce « patchwork », il est possible d'expliquer les grands principes de son fonctionnement dans une documentation accessible aux médecins et au grand public (Pôle qualité des données, *ibid*).

LA PRATIQUE DES SOINS À L'ÈRE DES ALGORITHMES

L'algorithme configure l'agency des acteurs

Avec l'introduction du score cœur, les médecins semblent avoir perdu l'autonomie dont ils disposaient dans le choix des patients à greffer. Ils doivent définir la stratégie de prise en charge en fonction des sorties de l'algorithme. Le nombre de points devient un argument pour justifier la stratégie thérapeutique lors des réunions d'équipe ou en entretien comme ici avec un cardiologue :

« si on veut envisager une greffe cardiaque, mais qu'on se rend compte [que le patient] a un score faible, donc un rang faible, on se dit qu'il ne peut pas avoir accès à une greffe, ça nous pousse à envisager d'autres alternatives thérapeutiques comme l'assistance ventriculaire [...] À l'inverse, si on sait que le patient est très haut, [...] on peut se laisser un peu de temps et ça change la prise en charge en fait »

Cette forme de gouvernance revêt ici un caractère plus autoritaire que les traditionnels indicateurs hospitaliers, dans la mesure où elle affecte plus directement les pratiques médicales (Bertillot, 2016). Parce qu'il est intégré dans le code informatique du système automatisé de décision, son déploiement dans tous les services est immédiat et inéluctable, à tel point que le nombre de points devient un argument à part entière dans l'élaboration de la stratégie de soin. Pour conserver une activité de greffe, il faut se plier aux règles et donc intégrer les indicateurs choisis par l'ABM qui ont un impact très pragmatique sur la prise en charge des malades.

Comme bien souvent, ces indicateurs ne sont pas neutres. Dans le cas du score cœur, le système charrie une représentation parmi d'autres du bénéfice attendu de la greffe, l'évitement du décès, qui n'est pourtant pas partagée par l'ensemble des médecins. Grâce à ce critère facilement quantifiable, l'ABM tire profit d'une prophétie auto-réalisatrice. Parmi les candidats qui ont des chances raisonnables de survivre à la transplantation, les plus graves sont choisis par le système national d'attribution. Cela réduit mécaniquement le nombre de décès ce qui est précisément l'objectif mesurable choisi par l'ABM. Imposer cette représentation permet à l'ABM de justifier plus facilement une approche centralisée de l'attribution des greffons que si le système s'était appuyé sur un critère moins facilement quantifiable, comme le bien-être des personnes greffées. Pourtant, comme nous allons le voir, cette approche quantifiée et visant prioritairement à éviter les décès n'est pas adoptée naturellement par les services de greffe.

À la différence de l'algorithme, les médecins des centres de greffe peuvent puiser dans une représentation plus large du patient et des bénéfices attendus de la greffe⁹. Cette représentation inclue les critères physiologiques pris en compte par le score pour estimer le risque de décès, mais d'autres critères peuvent entrer en jeu. Par exemple, la notion de patient *jeune* semble être un critère déterminant de la perception de la priorité. Pour un risque de décès équivalent, le médecin semble nettement plus investi dans le devenir d'un patient jeune. Bien qu'une préférence pour cette catégorie de patients soit prévue dans le score, c'est le sujet de beaucoup de situations de désaccord entre les choix des médecins et ceux de l'algorithme, comme celle présentée ci-dessous. D'autres critères moins quantifiables sont cités comme la

⁹ D'autres travaux ont déjà mis en lumière la dimension sociale (Loretti, 2019) ou personnelle (Dodier, 1991) de la représentation qu'ont les médecins de leurs patients, de leurs souffrances et des bénéfices qu'ils peuvent attendre d'une prescription médicale.

dynamique d'aggravation, la dénutrition, l'état d'anxiété, « *la perception d'un épuisement moral de l'attente* », ou « *la perception d'un idéal matching entre le donneur et le receveur* ». Une critique soulignée fréquemment par les médecins est la priorité systématique donnée aux patients dans un état de gravité extrême. Puisque la survie après la greffe ne peut pas être estimée correctement, la gravité prend une place prépondérante dans la valeur finale du score. Compte tenu de la rareté de la ressource, ces médecins déplorent que tous les patients doivent passer par une phase d'aggravation avant d'être traités, alors même que cela réduit les chances d'une greffe que les médecins jugeraient utile. Je rapporte ici les propos d'une cardiologue qui fait ce reproche de manière générale, puis s'appuie sur un exemple :

« Cardiologue : c'est le cas frontière du patient qui est grave, mais pas encore trop grave et on aimerait bien qu'il ait un score plus haut. Pour donner un exemple concret. On a eu un jeune, 37 ans, je crois, qu'on a reçu aux soins intensifs il y a deux semaines. Choc cardiogénique inaugural, donc il n'était ni connu, ni inscrit sur liste d'attente et ce patient est "insevrable" d'amine* donc on sait qu'il n'y aura pas d'alternative et que la greffe est la seule solution. Sauf qu'on lui prédisait un score pas très élevé, car il avait un bilan normal, car il était sous amines depuis des jours. Et le problème était son groupe « O » et 1 m 90 pour 85 kg donc c'est le genre de gabarit pour lesquels on n'a pas de greffons facilement. Donc on s'est dit qu'est-ce qu'on fait : "on va l'inscrire, il va avoir un score dans les 500-600. En plus c'est un gros gabarit donc on n'aura pas de greffons facilement". Puisqu'il était "insevrable", on lui a mis un heartmate [assistance ventriculaire longue durée], car on savait qu'on n'allait pas le greffer rapidement. [...]

Clément Henin : d'accord donc si vous aviez eu la main sur les greffons vous l'auriez mis...

Cardiologue : Super urgence ! Sûre ! Parce qu'il a 37 ans, parce qu'il est "insevrable" de dobutamine donc on sait qu'il n'y a pas d'autres solutions et on n'aurait pas forcément mis un heartmate d'emblée. C'est quand même quelque chose pour les résultats. En plus, il a fait

***** Il a besoin de ce médicament pour assurer sa survie (qui ne peut pas être sevré).

une défaillance forte donc c'est un heartmate compliqué avec une mortalité qui est majeure. »

Le reproche principal ici est la priorité donnée aux patients graves, mais l'interviewée insiste aussi sur la dynamique d'aggravation qui n'est pas prise en compte par l'algorithme. Lorsqu'elle dit que le patient est « *grave, mais pas encore trop grave* », elle estime que son état va se détériorer prochainement et elle préférerait le greffer immédiatement pour éviter cette aggravation et d'autres complications liées à l'assistance ventriculaire longue durée. Dans la suite, je note que le statut de *jeune* est cité en premier à deux reprises lors de l'argumentaire, ce qui lui accorde une importance particulière. Aussi, en signalant qu'il s'agit d'un choc inaugural, elle semble préciser que le patient, hormis cet épisode, est en bon état général. Alors que le score de l'ABM évite le décès des patients graves et renvoie implicitement les patients moins graves vers une alternative, la cardiologue aurait privilégié une transplantation rapide chez ce patient *jeune*, en bon état général et pour lequel elle prévoit une aggravation sous assistance ventriculaire.

Dans le sens d'autres études ethnographiques sur les algorithmes (Brayne et Christin, *ibid*), le savoir expérientiel propre à la profession médicale se trouve en tension avec le mode de connaissance basé sur les données propres à l'algorithme. Il est par ailleurs possible de donner une interprétation foucauldienne de cette destitution. D'abord, la création de la base de données Cristal, qui justifie la validité des décisions du score cœur, peut se rapporter à la production d'une vérité conduisant à l'établissement d'une politique de « *conduite des conduites* » qui incite les médecins à agir en fonction des points qui sont attribués à leurs patients. Plus particulièrement, une logique de « *direction des comportements* » (Badouard *et al.*, *ibid*) semble s'appliquer. En donnant des points aux patients les plus graves ou en pénalisant directement certaines pratiques, le score dirige le comportement des médecins vers le comportement voulu par l'ABM. Le score force un alignement d'intérêt entre les médecins qui veulent greffer et l'ABM qui souhaite greffer les patients les plus graves grâce au système de points.

Une décision trop importante pour le jugement humain ?

Bien que des praticiens, dont la cardiologue citée plus haut, déplorent que les décisions prises par le score soient moins bonnes que les décisions

humaines, le score est plutôt bien reçu par la communauté des transplantateurs. Compte tenu de l'historique des systèmes d'attribution rythmé par de nombreux abus, cette adhésion paraît en fait peu surprenante. Les efforts déployés par l'ABM pour rendre le score indépendant de la pratique médicale, même si comme nous le verrons plus loin il reste aux médecins une certaine marge de manœuvre, sont salués par les transplantateurs interrogés :

« Personnellement, j'ai bien aimé ce score, parce que c'est plus rationnel qu'avant. [...] on a perdu de la marge de manœuvre, mais cela a été introduit parce que la super urgence avait été abusée. Quand vous aviez un malade entre-deux, on lui mettait de la dobutamine et il finissait en haut de la liste puis greffé. »

Cette propriété se retrouve régulièrement sous les termes d'objectivité ou de neutralité de l'algorithme. La « subjectivité » du professionnel a eu dans les entretiens une connotation généralement négative. Il s'agit d'un élément qu'il faut limiter, voire écarter complètement. À titre d'exemple, je rapporte les propos d'un cardiologue, largement favorable au score, qui en défend l'objectivité :

« C'est l'avantage des scores et des choses comme ça, c'est d'être aussi objectifs que possible pour ne pas rentrer dans des variables trop subjectives et qu'on ne greffe pas des malades qui ne doivent pas être greffés par rapport à d'autres centres. Il faut que le clinicien ne puisse pas trop donner sa subjectivité ou alors que ce soient des experts vraiment indépendants. »

Un argument régulièrement cité est qu'il serait trop difficile pour les médecins qui sont au plus proche de la souffrance des patients de ne pas tout mettre en œuvre pour leur prodiguer le meilleur traitement possible même si cela défavorise les patients d'autres centres. Prenons pour exemple les propos d'un anesthésiste-réanimateur qui reconnaît être prêt à jouer avec les règles du score si la situation de son patient lui semble critique :

« Quand on pense que notre malade a besoin d'un nouveau cœur, ben voilà nous on a notre malade devant nous et on fait le maximum pour qu'il ait un cœur dans les limites de la légalité. Mais voilà s'il

faut faire certaines choses pour augmenter ses chances d'avoir un greffon on les fait. »

Il semble ici que les médecins soient mis en tension entre une préférence individuelle pour leurs malades, dont ils sont témoins directs des souffrances et une position d'extériorité par rapport aux jugements individuels visant à optimiser l'allocation d'une ressource rare. Alors que les deux positions sont moralement justifiables, la seconde est la priorité de l'action publique puisqu'elle est notamment garante de l'égalité d'accès au traitement. On comprend alors pourquoi, comme avaient fini par le faire les comptables décrits par Porter, les médecins sont obligés d'accepter un système qu'ils jugent pourtant moins performant et précis, mais globalement plus juste. L'adhésion des professionnels de santé à un système électronique bureaucratique au détriment de leur propre discrétion est en accord avec d'autres recherches dans le domaine (Reich, 2012). Selon Reich, les docteurs *disciplinés* continueraient à avoir un sentiment de contrôle de leur profession, à mesure que le professionnalisme lui-même deviendrait plus technocratique.

Compréhension et appropriation de l'algorithme

La compréhension du score par les médecins est loin d'être uniforme. Lors des entretiens, j'ai systématiquement demandé des explications sur l'algorithme ainsi que la liste des critères qu'il prend en compte afin d'estimer le niveau de connaissance des répondants. Il apparaît que certaines variables sont bien connues alors que d'autres ne sont presque jamais citées. La prise en compte de l'assistance veino-artérielle par exemple est connue de tous sauf une personne, alors que seules les deux secrétaires interrogées savent que les sexes du donneur et du receveur sont pris en compte. De la même manière, moins de 20 % des répondants savent que la maladie initiale est prise en compte. Une première explication est que les utilisateurs de l'algorithme se focalisent sur les variables les plus importantes du système. Cela est corroboré par certaines simulations grâce auxquelles j'ai pu mesurer l'importance des variables les mieux connues et par les propos de cette cardiologue : « *C'est principalement celles que je regarde parce que je sais que c'est là que sont les points. C'est ça qui va pondérer mon patient et le classer.* »

Par ailleurs, alors que le score est composé de quatre éléments principaux (risque de décès pré-greffe, risque de décès post-greffe, appariement physiologique et appariement géographique), l'effort de compréhension se focalise en quasi-totalité sur le risque de décès pré-greffe. Or, j'ai également pu mesurer que l'importance des variables d'appariement est au moins du même ordre que celle des variables de gravité. L'interprétation que je peux donner à ce fait est que les médecins focalisent leur attention là où ils ont un contrôle ou une impression de contrôle. Ils se désintéressent ainsi de l'appariement entre le donneur et le receveur, car cela ne constitue pas un levier qu'ils peuvent utiliser pour influencer les pratiques. Par exemple, ce cardiologue se montre très pragmatique dans le choix des variables dont il a connaissance en déclarant : « *Je connais les choses qui permettent de faire monter le score.* ». On voit donc que la compréhension du score s'intègre dans une dynamique de prise de contrôle par les médecins du score qui reste toutefois relative et très contrainte en comparaison de l'autonomie que les médecins avaient dans le système antérieur au score.

Une deuxième observation issue des entretiens est que seule une partie des médecins s'intéresse au fonctionnement du score. Environ la moitié des répondants connaît moins d'un tiers des critères pris en compte par le score et beaucoup déclarent avoir des *lacunes* sur son fonctionnement. Certains montrent même un grand désintérêt pour le fonctionnement de l'algorithme et préfèrent se reposer sur leurs *collègues* plus *investis* dans le système. Dans les quatre centres interrogés, seules quelques personnes sont investies dans le score et jouent un rôle de référent. Il s'agit le plus souvent de cardiologues non jeunes jouissant d'un certain niveau dans la hiérarchie. En raison de sa complexité, le coût d'apprentissage ne se justifie que par des interactions régulières avec le score ce qui conduit à une distribution inéquitable de la connaissance au sein des équipes. Par ailleurs, ceux qui concentrent la connaissance se retrouvent aussi en première ligne pour propager la confiance ou la défiance en l'algorithme au sein des équipes. Rejoignant les résultats de (Elish, 2018), on voit bien ici que les notions de confiance, de preuve et d'autorité de l'algorithme s'établissent plus par des réseaux sociaux existants que par des propriétés techniques de l'algorithme. En outre, ce référent est responsable de la prise en compte du score dans l'établissement de la stratégie de soin et il est aussi responsable de la participation du centre aux évolutions du système, deux aspects qui sont traités dans les parties suivantes.

Le déploiement des « tactiques »

Alors que le projet du score cœur retire aux médecins toute autonomie dans l'attribution des greffons, limitant leur pouvoir discrétionnaire à l'inscription et à l'acceptation de la proposition, l'observation de terrain montre qu'au travers de certains usages, les médecins parviennent à influencer le score. Ces usages peuvent être vus comme des « tactiques » au sens de De Certeau (De Certeau, 1994), c'est-à-dire les pratiques visant à contourner ou déjouer un contrôle imposé de l'extérieur en exploitant ses failles. Les médecins ont appris à jouer avec la norme pour *optimiser* le score des patients qu'ils estiment les plus prioritaires. La cardiologue dont je rapporte les propos déplore la mise en place de cette *machine perverse* : « *C'est un petit peu une machine perverse [...] dans la façon dont on aborde les choses. En fait on cherche à grappiller des points parce qu'on veut rendre service à notre patient, qui est certes plus grave donc il est légitime d'optimiser son score.* »

Ces tactiques sont des réponses classiques d'adaptation que l'on retrouve dans d'autres terrains (Brayne et Christin, *ibid* ; Cardon, *ibid* ; Christin, *ibid*). Un glissement s'est manifestement opéré entre le projet d'un outil de mesure, visant à estimer objectivement la priorité des candidats à la greffe et des points que l'on *cherche à grappiller*. Ce glissement est décrit par ailleurs comme *pervers*, car les médecins sont conscients que l'algorithme défend un intérêt commun d'équité. Prenons quelques exemples pour éclairer ce propos.

Malgré les efforts déployés par l'ABM pour rendre le score indépendant des pratiques médicales, certains traitements, dont l'utilisation est laissée à la discrétion des médecins, sont pris en compte dans la qualification de la gravité. Parmi les arguments employés pour justifier leur utilisation, il existe une zone grise entre une motivation clinique concernant les effets du traitement sur le corps du malade et une manœuvre tactique visant à « *optimiser* » le score du candidat. Par exemple, lors d'une réunion d'équipe à laquelle j'ai assisté, la mise en place d'une assistance veino-artérielle a été suggérée pour un patient¹⁰. Comme l'indication pour ce traitement n'était pas

¹⁰ L'assistance véino-artérielle, qui consiste à remplacer temporairement le travail cœur à l'aide d'une pompe extérieure, nécessite une intervention chirurgicale qui s'accompagne d'un risque important de complications (lésion vasculaire, thrombose, infection). Il s'agit d'une intervention d'urgence pour assurer une circulation sanguine minimale nécessaire au

formelle, la discussion a porté sur les points qu'elle pourrait rapporter. Puisque le score du patient était déjà très haut, l'équipe a décidé de rejeter l'utilisation de cette assistance. Plus généralement, ce médecin reconnaît que cette assistance (dont l'abréviation anglaise est ECMO) risque d'être utilisée de plus en plus fréquemment pour manipuler le système :

« Maintenant c'est l'ECMO qui nous donne 900 points [...]. Parfois on a envie de poser l'ECMO, d'avoir 900 points d'emblée et de greffer rapidement le patient. [...] C'est ça l'aberration du système. Avant on avait la dobutamine, maintenant on a l'ECMO. »

Cet extrait dévoile un jeu de chat et de souris auquel se livrent les médecins et l'ABM. Pour limiter les possibilités de manipulation du score, l'ABM construit un score qui ne prend pas en compte la dobutamine. Progressivement, les médecins s'adaptent au nouveau score en employant une ECMO pour gagner des points. Cette mécanique d'adaptation rappelle les évolutions de l'algorithme PageRank de Google pour s'adapter aux pratiques des Webmestres (Cardon, *ibid*). Dans le cas du score cœur, en passant d'un médicament (dobutamine) à un traitement nécessitant une chirurgie et comportant un risque plus important pour le patient (ECMO), l'ABM mise sur la bienveillance des médecins peu disposés à risquer la santé de leurs patients pour gagner des points. Si cette stratégie semble relativement efficace contre les médecins, elle protège moins bien contre l'optimisation par le patient lui-même qui a moins de problèmes éthiques à dégrader son propre état de santé. Au moins un patient, qui est par ailleurs professionnel de santé, a refusé la pose d'une assistance ventriculaire longue durée dans le but d'accéder plus rapidement à la greffe en pariant sur une dégradation de son état¹¹. L'importance de la compréhension de l'algorithme dans le déploiement de tactiques apparaît ici nettement, car les connaissances médicales dont disposait le patient ont été nécessaires à sa mise en place.

Un autre mécanisme peut aussi être utilisé par les médecins pour influencer le score cœur : les composantes experts. Il s'agit d'un système d'exception qui accorde un grand nombre de points pour des situations cliniques spécifiques dont la liste est établie par l'ABM. Les demandes de composantes sont évaluées par un expert extérieur au centre hospitalier.

fonctionnement des autres organes vitaux.

¹¹ La priorité d'accès aux patients sous assistance longue durée a été revue à la hausse depuis l'écriture de cet article et le calcul très risqué fait par ce patient n'a plus vraiment lieu d'être aujourd'hui.

Puisque c'est l'expert qui tranche, il existe une certaine marge d'interprétation. Ces composantes sont vues comme *une carte à jouer* lorsque les médecins jugent que la gravité d'une situation ne se reflète pas dans le risque prédit. Par exemple, dans la situation décrite ci-dessous, la cardiologue « *demande quand même* » une composante expert, même si elle sait que la patiente est « *un petit peu hors schéma* » en expliquant le dossier à l'expert :

« [...] quand vraiment je vois que le patient s'aggrave, je demande quand même une composante expert en expliquant qu'on est un petit peu hors schéma. J'explique le dossier. J'ai eu par exemple le cas pour une jeune patiente qui était en attente de retransplantation, qui a 20 ans et qui a été greffée à l'âge de deux ans environ. [...] j'ai pris la composante expert patiente non 'assistable', mais c'est un profil qui est loin d'être typique si vous voulez. »

Elle estime que la gravité réelle ne se reflète pas dans le score de sa patiente et espère que l'expert sera du même avis et lui accordera la composante, bien que cette situation ne soit pas explicitement prévue par le système. Enfin, les évolutions de l'état de gravité du patient peuvent aussi être mises à profit pour « *optimiser* » le score. Des bilans doivent être transmis régulièrement à l'ABM pour mettre à jour le score. En cas d'aggravation, les médecins peuvent faire un bilan prématuré et mettre à jour la valeur du score, le patient apparaît ainsi dans l'état le plus grave. À l'inverse, si le bilan reflète une amélioration, le médecin peut attendre la fin du délai réglementaire.

Les réunions du groupe technique restreint score cœur

En plus des tactiques que les médecins peuvent déployer ponctuellement, des réunions techniques sont organisées régulièrement à l'ABM pour que les centres puissent discuter de l'évolution de la stratégie du score. Christian Jacquelinet, l'un des concepteurs du score, a longuement insisté sur l'importance de ces réunions pour l'adhésion des centres au système et pour assurer que le score s'adapte au mieux à la réalité de la prise en charge. Nous décrivons ici par deux exemples le fonctionnement de ces réunions.

Le premier exemple est raconté par une cardiologue qui a détecté, peu de temps après l'introduction du score, un défaut dans les règles d'attribution pour les patients jeunes :

« Je suis allée à une réunion, j'étais un peu effarée. On avait eu un jeune monsieur qui attendait un greffon, 25 ans, je crois, avec beaucoup de points. Il devait être en 7e ou 8e position et devant lui il y avait des patients de 65 ans et tout ça ! On a eu de la chance, car devant nous tout le monde a refusé. Je me suis rendu compte que quelque chose m'avait échappé, j'en ai parlé au niveau de l'agence quand on a eu une réunion. Tout le monde a dit : "c'est pas normal". [...] On a fait changer les règles pour que les jeunes ne perdent pas de points quand il y a une différence d'âge. [...] en fait c'est en regardant ça que j'avais interpellé à la réunion. J'avais même fait une copie d'écran et j'étais allée avec ma copie d'écran à l'ABM et je leur avais dit : "ben regardez !". »

C'est en regardant le classement des autres candidats que cette cardiologue se rend compte que des patients âgés pouvaient être prioritaires sur son patient jeune, ce qui lui a paru inacceptable. Son point de vue est partagé par l'ensemble des centres de greffe et par l'ABM et les règles sont modifiées en conséquence. Dans cet exemple, l'appropriation du système et l'espace de discussion mis en place par l'ABM permettent aux professionnels d'influencer le score de manière permanente.

La réunion de novembre 2019 a concerné les patients en attente de greffe sous assistance ventriculaire longue durée¹². Puisque l'assistance améliore la condition générale des patients et que le score est basé sur la gravité, ces derniers ont tendance à être « oubliés par le score » et à n'être greffés qu'à la suite d'une aggravation significative. La prise en compte des assistances ventriculaires dans le score a suscité de nombreux débats. Une cardiologue déclare avoir mentionné ce problème avant la mise en place du score dès 2017, mais les arguments des centres sont difficilement entendus. Dans la représentation de l'ABM où l'évitement du décès occupe le rôle principal, prendre en compte les « oubliés » sous assistance ventriculaire longue durée

¹² L'assistance ventriculaire longue durée est un appareil mécanique implanté sur le cœur qui l'assiste dans sa fonction pompe assurant un débit cardiaque nécessaire à une bonne circulation sanguine. À la différence de l'assistance veino-artérielle, cette assistance a une durée de plusieurs mois et est compatible avec un retour au domicile.

n'améliore pas les résultats de la greffe, car le taux de survie mesuré est similaire que le patient soit greffé avant ou après une aggravation. À l'inverse, les arguments des centres promoteurs de ce changement considèrent le parcours de soin et la qualité de vie des patients, trouvant regrettable de devoir attendre de se dégrader pour accéder au traitement. Finalement, la priorité des patients sous assistance longue durée est revue à la hausse.

Ces exemples montrent la souplesse relative de l'ABM sur l'évolution des règles d'attribution. L'avis des centres est pris en compte, mais plus difficilement lorsqu'il ne va pas dans le sens d'une attribution basée sur l'évitement du décès. Plus généralement, le rôle ces réunions met en avant l'importance de l'appropriation par les médecins des traitements effectués par l'algorithme pour permettre la prise en compte de leur avis par l'ABM. Cette compréhension est nécessaire pour déconstruire les choix implicites faits lors de l'élaboration. Ces réunions peuvent être vues comme l'une des formes de « *buffering* » proposée par Angèle Christin (Christin, *ibid*) : l'« *open-critique* ». Dans ce cas, la critique est à la fois informée et écoutée, car elle peut conduire à des évolutions concrètes du système d'attribution national.

CONCLUSION : UN APPORT À « L'EXPLAINABLE AI » (XAI)

La question de l'opacité des algorithmes, notamment ceux basés sur l'intelligence artificielle, reçoit une attention grandissante de la part des universitaires dans les domaines de l'informatique avec notamment un appel d'une agence militaire de recherche américaine (Gunning, 2017), du droit avec l'introduction dans la réglementation européenne sur les données personnelles d'un droit à l'explication (Goodman et Flaxman, 2017) ou la publication d'ouvrages sur la question des décisions automatiques (Desmoulin-Canselier et Le Métayer, 2020) et des sciences sociales (Kellogg *et al.*, 2020). En informatique, le défi principal est l'élaboration d'outils techniques expliquant aux concepteurs et aux utilisateurs le fonctionnement d'algorithmes ou explicitant les raisonnements ayant conduit à des décisions particulières (Mittelstadt *et al.*, 2018). À la suite notamment d'un appel de Tim Miller en 2017 (Miller *et al.*, 2017), l'intérêt des sciences sociales pour la conception de ces explications est reconnu par une partie de la communauté (Miller, 2017 ; Mueller *et al.*, 2019), notamment pour identifier

les mécanismes à l'œuvre dans les processus d'explication humains. Cela a par exemple conduit au développement de méthodes d'explication interactives (Henin et Le Métayer, 2020b ; 2021) et à des méthodes de justifications d'algorithmes (Henin et Le Métayer, 2020a).

L'analyse menée sur le cas d'un algorithme utilisé à l'hôpital offre un nouveau regard au besoin de compréhension et d'explication. En premier lieu, il apparaît que la notion de confiance en l'algorithme, qui est pourtant une des principales motivations de la littérature XAI, ne semble pas trouver sa réponse dans une compréhension plus fine du système, mais plutôt dans une observation de ses effets et de ses résultats. La critique n'apparaît pas à la lecture du code, mais plutôt dans l'observation des décisions prises par l'algorithme. De la même manière, la compréhension de l'algorithme ne semble pas jouer un rôle significatif sur l'acceptation du système, ce qui est ici encore à contre-courant d'idées communes dans la communauté XAI. Comme pour les normes en cancérologie étudiées par Castel et Merle (Castel et Merle, 2002), l'acceptation du système semble plutôt conditionnée à la possibilité qu'ont les acteurs à la fois de conserver une certaine marge de manœuvre et de garder une voix dans l'élaboration du système.

Alors que favoriser la confiance dans les algorithmes ou faciliter leur acceptation par les professionnels semble être le but global de la mouvance XAI, notre terrain nous montre que les médecins n'ont pas besoin de comprendre précisément le fonctionnement système pour lui accorder une certaine confiance et l'accepter. Comme pour d'autres systèmes dans la vie courante, il semble que l'algorithme peut conserver le statut de boîte noire et que son comportement peut être décortiqué, analysé sans entrer dans le détail de son fonctionnement. En effet, alors qu'aucun des médecins interrogés n'a de compétence particulière en informatique ou ne s'est intéressé au fonctionnement algorithmique du score, ces derniers parviennent très bien à jouer de ses règles et à participer à ses évolutions, de la même manière que le conducteur d'une voiture parvient à contrôler son véhicule sans forcément comprendre le fonctionnement du moteur. Ainsi, le champ de recherche XAI devrait focaliser ses efforts sur l'élaboration de méthodes permettant non pas d'expliquer le fonctionnement du système, mais de montrer son comportement d'utilisation ainsi que les leviers pouvant être utilisés pour se l'approprier.

RÉFÉRENCES

AGENCE DE LA BIOMEDECINE, (2016), Site de l'Agence de la Biomédecine, [En ligne] Disponible à l'adresse : <https://www.agence-biomedecine.fr/annexes/bilan2016/donnees/organes/03-coeur/synthese.htm> (consulté le 25/11/20).

AGENCE DE LA BIOMEDECINE, (2017), Agence de la biomédecine - Le rapport annuel médical et scientifique 2017, [En ligne] Disponible à l'adresse : <https://www.agence-biomedecine.fr/annexes/bilan2017/donnees/organes/03-coeur/synthese.htm> (consulté le 11/08/20).

AGENCE DE LA BIOMEDECINE, (2019), Rapport annuel 2018, Disponible à l'adresse suivante : <https://donsdegametes-solidaires.fr/wp-content/uploads/2019/09/rapport-annuel-2018.pdf> (consulté le 25/11/20)

BADOUARD R., MABI C., SIRE G. (2016), Beyond « Points of Control »: logics of digital governmentality, *Internet Policy Review*, vol. 5, n°3, [En ligne] Disponible à l'adresse : <https://policyreview.info/articles/analysis/beyond-points-control-logics-digital-governmentality> (consulté le 25/11/2020).

BENBOUZID B., (2018), Quand prédire, c'est gérer : La police prédictive aux États-Unis, *Réseaux*, vol. n° 211, n°5, p. 221-256.

BERTILLOT H., (2016), Des indicateurs pour gouverner la qualité hospitalière. Sociogenèse d'une rationalisation en douceur, *Sociologie du Travail*, vol. 58, n°3, p. 227-252.

BRAYNE S., CHRISTIN A. (2020), Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts, *Social Problems*.

BURRELL J. (2016), How the machine « thinks »: Understanding opacity in machine learning algorithms, *Big Data & Society*, vol. 3, n°1.

CALABRESI G., BOBBITT P. (1978), *Tragic choices*, New York, NY, W.W. Norton & Company.

CANTRELLE C. (2018), « Le parcours de soin des greffés cardiaques en France : détermination des facteurs associés à leur accès à la greffe » Thèse de doctorat : Santé publique, Université Paris-saclay.

CARDON D. (2013), Dans l'esprit du PageRank : Une enquête sur l'algorithme de Google, *Réseaux*, vol. n° 177, n°1, p. 63-95.

CASTEL P., MERLE I. (2002), Quand les normes de pratiques deviennent une ressource pour les médecins, *Sociologie du travail*, p. 337-355.

CHILDRESS J. F. (2001), Putting Patients First in Organ Allocation: An Ethical Analysis of the U.S. Debate, *Cambridge Quarterly of Healthcare Ethics*, vol. 10, n°4, p. 365-376.

CHRISTIN A. (2017), Algorithms in practice: Comparing web journalism and criminal justice, *Big Data & Society*, vol. 4, n°2.

COQUIDE P. (1993), Le marché sauvage des greffes d'organes, L'EXPRESS, 9 décembre, [En ligne] Disponible à l'adresse : https://lexpansion.lexpress.fr/actualite-economique/le-marche-sauvage-des-greffes-d-organes_1383766.html (consulté le 06/08/20)

DE CERTEAU M. (1980), *L'invention du quotidien*, Paris, Gallimard.

DECHAMP-LE ROUX C. (1997), Débats autour de la transplantation d'organes, *Sciences sociales et santé*, vol. 15, n°1, p. 99-127.

DESMOULIN-CANSELIER, S. & LE MÉTAYER, D., (2020), *Décider avec les algorithmes. Quelle place pour l'homme, quelle place pour le droit ?*, Dalloz, Paris.

DESROSIERES A. (2010), *La politique des grands nombres. Histoire de la raison statistique*, La Découverte, Paris.

DODIER N. (1991), Expérience privée des personnes et expertises médico-administratives. Une enquête dans la médecine du travail., *Sciences sociales et santé*, vol. 9, n°2, p. 79-121.

DORENT R., CANTRELLE C., JASSERON C., LEGEAI C. (2014), La greffe cardiaque en France : bilan actuel, *La Presse Médicale*, vol. 43, n°7-8, p. 813-822.

DORENT R., EPAILLY E., SEBBAG L., (2011), The effect of graft allocation system on outcomes in heart transplantation in France: Has the time come to take calculated survival benefit into account?, *The Journal of Heart and Lung Transplantation*, vol. 30, n°11, p. 1299-1300.

EGAN T. M. (2017), Cardiac Risk Score—Halfway There, *Transplantation*, vol. 101, n° 12, p. 350.

ELISH M. C. (2018), The Stakes of Uncertainty: Developing and Integrating Machine Learning in Clinical Care, *Ethnographic Praxis in Industry Conference Proceedings*, vol. 2018, n°1, p. 364-380.

ELWYN G., SCHOLL I., TIETBOHL C., MANN M., EDWARDS A. G., CLAY C., FROSCH D. L. (2013), « Many miles to go ... »: a systematic review of the implementation of patient decision support interventions into routine clinical practice, *BMC Medical Informatics and Decision Making*, vol. 13, n° S2.

GOODMAN B., FLAXMAN S. (2017), European Union regulations on algorithmic decision-making and a « right to explanation », *AI Magazine*, vol. 38, n° 3, p. 50.

GUNNING D., (2017), Explainable Artificial Intelligence (XAI), [En ligne] Disponible à l'adresse : [https://www.cc.gatech.edu/~alanwags/DLAI2016/\(Gunning\)%20IJCAI-16%20DLAI%20WS.pdf](https://www.cc.gatech.edu/~alanwags/DLAI2016/(Gunning)%20IJCAI-16%20DLAI%20WS.pdf) (consulté le 25/11/20).

HENIN, C. & LE MÉTAYER, D., (2020a), Towards a framework for challenging ML-based decisions, Présenté à : *1st International Workshop on Deceptive AI@ ECAI2020*.

HENIN, C. & LE MÉTAYER, D., (2020b), A generic framework for black-box explanations, Présenté à : International Workshop on Fair and Interpretable Learning Algorithms (FILA 2020), IEEE.

HENIN, C. & LE MÉTAYER, D., (2021), A Multi-layered Approach for Tailored Black-box Explanations, Présenté à : ICPR'2020 WorkShop Explainable Deep Learning-AI, Springer.

JASSERON C., LEGEAI C., JACQUELINET C., LEPRINCE P., CANTRELLE C., AUDRY B., DORENT R., (2017), Prediction of Waitlist Mortality in Adult Heart Transplant Candidates: The Candidate Risk Score, *The Journal of Heart and Lung Transplantation*, vol. 36, n° 4, p. S114.

JASSERON C., LEGEAI C., JACQUELINET C., NUBRET-LE CONIAT K., FLECHER E., CANTRELLE C., DORENT R., (2018), Optimization of heart allocation: The transplant risk score, *American Journal of Transplantation*, vol. 19, n° 5, p. 1507-1517.

KELLOGG K. C., VALENTINE M. A., CHRISTIN A. (2020), Algorithms at Work: The New Contested Terrain of Control, *Academy of Management Annals*, vol. 14, n° 1, p. 366-410.

LORETTI A. (2019), Les logiques sociales de la décision médicale. Étude des critères de prescription médicale en cancérologie des voies aérodigestives supérieures, *Sciences sociales et santé*, vol. 37, n° 4, p. 73.

MILLER T. (2017), Explanation in Artificial Intelligence: Insights from the Social Sciences, *Artificial Intelligence*, vol. 267.

MILLER T., HOWE P., SONENBERG L., (2017), Explainable AI: Beware of inmates running the asylum, *In: IJCAI-17 Workshop on Explainable AI (XAI)* vol. 36.

MITTELSTADT, B., RUSSEL, C. & WACHTER, S., (2019), Explaining explanations in AI, *In : Proceedings of the conference on fairness, accountability, and transparency*, p. 279–288.

MUELLER S. T., HOFFMAN R. R., CLANCEY W., EMREY A., (2019), Explanation in Human-AI Systems: A Literature Meta-Review Synopsis of Key Ideas and Publications and Bibliography for Explainable AI, Technical Report [En ligne] Disponible à l'adresse : <https://apps.dtic.mil/sti/pdfs/AD1073994.pdf> (consulté le 25/11/20)

PASQUALE F. (2015), *The black box society*, Harvard University Press.

PERSAD G., WERTHEIMER A., EMANUEL E. J. (2009), Principles for allocation of scarce medical interventions, *The Lancet*, vol. 373, n° 9661, p. 423-431.

POLE QUALITE DES DONNEES (2019), Guide du Score Cœur (no version v3) [En ligne] Disponible à l'adresse : https://www.agence-biomedecine.fr/IMG/pdf/guide_score_coeur_v3.pdf (consulté le 25/11/20).

PORTER T. M. (1996), *Trust in numbers: The pursuit of objectivity in science and public life*, Princeton University Press.

POTENA L., KHUSH K. K. (2017), The Conundrum of Equitable Organ Allocation in Heart Transplantation: The Moving Target of Candidate Risk Score, *Transplantation*, vol. 101, n° 9, p. 1969-1970.

REICH A. (2012), Disciplined doctors: The electronic medical record and physicians' changing relationship to medical knowledge, *Social Science & Medicine*, vol. 74, n° 7, p. 1021-1028.

ROTH A. E., SÖNMEZ T., ÜNVER M. U. (2004), Kidney exchange, *The Quarterly journal of economics*, vol. 119, n° 2, p. 457-488.

ROUVROY A., BERNS T. (2013), Gouvernamentalité algorithmique et perspectives d'émancipation : le disparate comme condition d'individuation par la relation ?, *Réseaux*, n° 177, n°1, p. 163-196.

STEINER P. (2010), *La transplantation d'organes : un commerce nouveau entre les êtres humains*, Gallimard, Paris.

STEINER P., (2017), Economy as matching, *Política & Sociedade*, vol. 18, n° 43, p. 14-45.

STEINER P., « *De corps à corps : l'appariement dans la transplantation rénale* », in SIMIONI M. et STEINER P. (Dir.), *Faire des paires. Sociologie de l'appariement*, (à paraître).

SUPIOT A. (2015), *La gouvernance par les nombres*, Fayard, Paris.

THIBAUT P. (2013), Allemagne : scandale de fraude au don d'organes, RFI, 20 août, [En ligne] Disponible à l'adresse : <https://www.rfi.fr/europe/20130820-allemande-scandale-fraude-don-organes-gottingen> (consulté le 25/11/2020)

VILLANI C. (2018), Donner un sens à l'intelligence artificielle, Rapport de la mission parlementaire, Ministère de l'enseignement supérieur de la recherche et de l'innovation, [En ligne] Disponible à l'adresse : https://www.aiforhumanity.fr/pdfs/9782111457089_Rapport_Villani_accessible.pdf (consulté le 25/11/2020).

YANG Q., ZIMMERMAN J., STEINFELD A., CAREY L., ANTAKI J. F. (2016), Investigating the Heart Pump Implant Decision Process: Opportunities for Decision Support Tools to Help, In : *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, ACM Press, Santa Clara, California, USA, p. 4477-4488.