



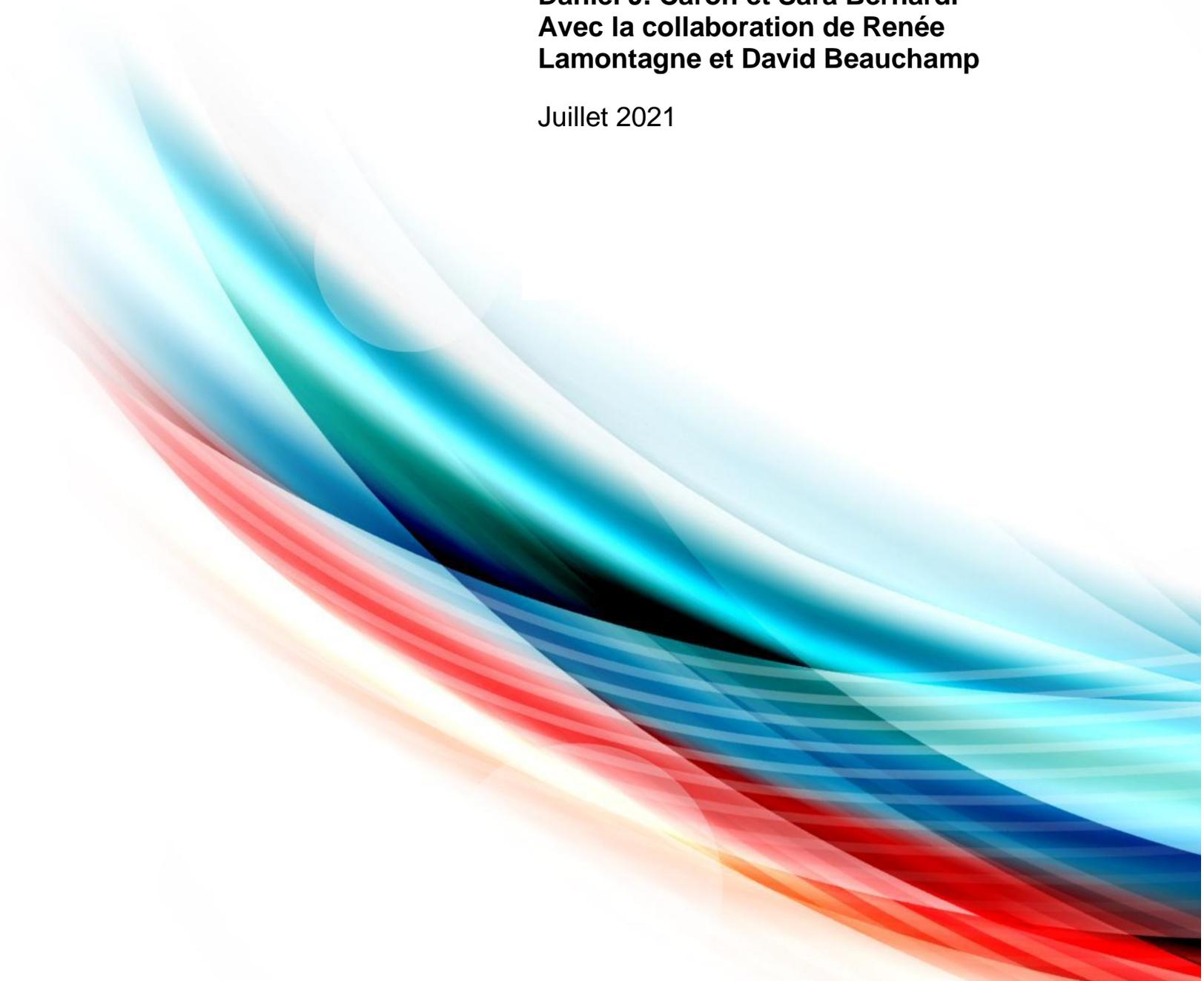
Chaire de recherche en exploitation des ressources informationnelles

 École nationale
d'administration publique

**Le partage des renseignements de santé et la performance du
système de santé : paramètres et enjeux**

**Daniel J. Caron et Sara Bernardi
Avec la collaboration de Renée
Lamontagne et David Beauchamp**

Juillet 2021



La Chaire de recherche en exploitation des ressources informationnelles de l'ENAP a pour mission l'étude, le questionnement et le développement des instruments de gouvernance nécessaires pour maximiser l'exploitation des ressources informationnelles dans l'univers numérique. Ces instruments sont principalement représentés par les politiques publiques, les politiques et pratiques administratives, la culture des organisations et les solutions technologiques. Les activités de recherche seront concentrées autour de l'étude de solutions pour permettre l'intégration des nouveaux comportements et des nouvelles technologies numériques afin d'assurer l'exploitation optimale des ressources informationnelles dans les organisations et la société de façon générale.

ISBN : 978-2-89734-067-4 (PDF)

Dépôt légal Bibliothèque et Archives nationales du Québec, 2021

Bibliothèque et Archives Canada, 2021

Table des matières

Introduction	3
Méthodologie	4
Les renseignements de santé, la recherche en santé et leur raison d'être respective	5
Raison d'être et objectifs pour la création des renseignements de santé : usage primaire	5
Une question centrale Objectifs de la recherche en santé	6
Un équilibre à atteindre entre risques et opportunités	8
Partage des renseignements de santé pour des fins de recherche	9
État de situation	9
Défis et enjeux	10
<i>L'accès aux données de santé</i>	10
<i>Les caractéristiques des données de santé</i>	11
<i>La protection du citoyen</i>	12
<i>La confiance et l'acceptabilité sociale</i>	15
Modèles recensés de partage des renseignements de santé	16
Présentation des modèles	16
<i>Le registre national</i>	16
<i>Les données ouvertes</i>	17
<i>Le modèle fiduciaire ou le tiers de confiance</i>	19
<i>Les banques de données liées (en anglais linked databank)</i>	20
<i>Les partenariats</i>	22
<i>Considérations pour un modèle de partage des données de santé avec les chercheurs</i>	23
Discussion et conclusion	24
Références	28
Annexe 1 Liste des mots-clés	35

Introduction

Les renseignements de santé et de services sociaux font partie intégrante du fonctionnement des systèmes de santé. Chaque interaction de l'utilisateur avec un professionnel de la santé et des services sociaux ou avec les organisations du réseau génère des données tant cliniques qu'administratives. Ces données ont toujours eu une grande valeur. Toutefois, le volume de données produites et surtout les capacités récentes de leur traitement se sont accrues substantiellement et font croître leur valeur à un rythme exponentiel (Ballantyne, 2018). Il faut rappeler que ces données sont d'abord produites pour répondre à un besoin primaire soit la prestation de services à l'utilisateur et la bonne gestion du système de santé et des services sociaux. Avec l'accroissement des capacités de traitement issu du développement rapide des technologies de l'information et des communications, ces données renferment d'autres potentiels d'utilisation. Par exemple, l'implantation au cours des dernières décennies de dossiers de santé électronique et l'avènement des technologies fondées sur les données ont grandement contribué à cet essor (Platt et al., 2019 ; Peugeot, 2018). On peut, par exemple, penser aux objets connectés qui produisent une très grande quantité de données. Ces usages secondaires sont multiples et de plus en plus d'acteurs convoitent les données afin de les valoriser dans d'autres contextes.

La recherche constitue l'un de ces usages secondaires. En effet, le système de soins de santé et sa performance s'appuient fortement sur la recherche pour améliorer ses pratiques tant du côté clinique qu'administratif. La disponibilité de données de haute qualité est donc essentielle à la recherche (Council of Canadian Academies, 2015) et les connaissances acquises sont fondamentales pour améliorer la santé en général et maintenir des soins de santé et des services sociaux de haute qualité. Bien qu'il existe de multiples bénéfices, l'utilisation des données de santé à des fins de recherche engendre des enjeux et des risques qui doivent être pris en considération dans le développement des pratiques et des politiques publiques et administratives à cet égard. Il faut donc avancer avec enthousiasme quant aux retombées possibles de cette utilisation, mais avec prudence (Peugeot, 2018, p. 30).

C'est dans cette optique que le présent rapport de recherche fondé principalement sur une revue de littérature et une analyse de ses principales constatations s'intéresse au partage des

renseignements de santé¹ à des fins de recherche. Par conséquent, il n’y sera pas fait état nommément du partage des renseignements liés à la prestation des services sociaux lorsque ceux-ci ne sont pas associés à la prestation de soins de santé – ou en tant que déterminants de la santé – ni des défis et des enjeux propres qui y sont liés. Toutefois, plusieurs défis et enjeux relatifs au partage des renseignements de santé concernent également la prestation des services sociaux, par exemple, la protection de la vie privée.

Après une description de la méthodologie utilisée, le rapport élabore sur l’enjeu des renseignements de santé dans un contexte de recherche en santé. Par la suite, le partage des renseignements de santé à des fins de recherche est abordé en présentant un état de situation et les enjeux et défis qui découlent des pratiques courantes. Le rapport présente une recension des principaux modèles de partage des renseignements de santé à des fins de recherche et conclut par une discussion et la présentation d’un « idéal type » de ce que pourrait être un modèle de partage en tant que levier de la performance du système de santé.

Méthodologie

Les propos du présent rapport sont issus d’une vaste revue de la littérature scientifique et grise portant sur le thème du partage des renseignements de santé. Au total, sept bases de données scientifiques ont été interrogées (Academic Search Complete, CAIRN, Emerald Journals, Elsevier Science Direct, Taylor and Francis, SAGE et Wiley). La recherche a également été étendue à Google Scholar afin d’assurer une plus grande exhaustivité des résultats. Enfin, pour obtenir de l’information concernant des exemples appliqués (cas concrets), des recherches sur des sites Internet gouvernementaux ont été lancées ainsi que sur ceux d’organismes reconnus comme l’Organisation de coopération et de développement économiques (OCDE).

Afin de baliser la recherche, sauf à quelques exceptions près, elle a été restreinte à la documentation produite après 2010. Les bases de données ont été consultées entre le 1er mai

¹ Bien que la terminologie « renseignements de santé » soit préconisée par le Ministère de la Santé et des Services sociaux, le terme « données de santé » sera également utilisé de manière interchangeable au sein de ce rapport, car il s’agit de la nomenclature principalement utilisée dans la littérature scientifique. De même, plusieurs recherches parlent de données administratives alors qu’elles incluent aussi des données cliniques. Dans le cadre de ce rapport-ci, nous utiliserons les termes données de santé de manière inclusive. Toutefois, des distinctions entre données cliniques et données administratives doivent être faites et elles seront discutées dans un rapport ultérieur intitulé *La gestion des renseignements de santé dans un contexte de gouvernance informationnelle responsable, juillet 2021*.

2021 et le 30 juin 2021. De plus, pour assurer une revue systématique des bases de données, une liste de mots-clés a été dressée en français et en anglais. Cette dernière est présentée à l'annexe 1. Au total plus de 345 articles ont été recensés. Après une première itération basée sur la pertinence, 52 sont utilisés dans la présente étude.

Les renseignements de santé, la recherche en santé et leur raison d'être respective

D'entrée de jeu, il est crucial de rappeler que les données de santé ne sont pas créées a priori pour des fins de recherche. Elles le sont d'abord et avant tout pour servir la prestation de soins aux usagers ainsi que pour effectuer les diverses redditions de comptes. Le contexte de réutilisation massive est assez nouveau et propre à l'univers numérique où les données connaissent une multitude d'usages secondaires. Ceci exige de porter une attention particulière afin de favoriser l'adéquation entre les usages primaires- prestation de soins et reddition de comptes – et les usages secondaires – recherche, innovation et commercialisation. Cette situation est d'autant plus complexe que les deux sphères sont interdépendantes. En effet, les résultats de la recherche, de l'innovation et de la commercialisation de certains produits à laquelle ils peuvent mener sont aussi des variables qui peuvent favoriser une meilleure performance du système de santé tant clinique qu'administrative.

Raison d'être et objectifs pour la création des renseignements de santé : usage primaire

Les soins de santé et de services sociaux sont reconnus comme faisant partie d'un système complexe, dans lequel se dessinent des interactions non linéaires entre les personnes, les processus et les technologies qui le composent (Lipsitz, 2012). De nombreux acteurs et processus médicaux sont imbriqués dans l'écosystème des soins de santé d'une seule personne (Spruit et Lytras, 2018 ; Caron, 2021a). Ces nombreuses interactions donnent lieu à la création d'une multitude de données afin de répondre à la production des divers extrants du système tant du côté clinique qu'administratif. Celles-ci contiennent des informations relatives à la santé actuelle et passée, aux conditions médicales, aux tests médicaux de l'utilisateur (Elger et al., 2010) ainsi que des données décrivant la trajectoire de soins, les délais, les dépenses, etc. Dans le contexte canadien, les données de santé comportent au moins trois caractéristiques uniques par rapport

aux données détenues par la plupart des autres pays (Quan et al., 2012). Tout d'abord, des registres d'assurance maladie sont tenus, ce qui constitue une importante base de données à grande échelle (Quan et al., 2012). Deuxièmement, la couverture complète des services hospitaliers et ambulatoires engendre des systèmes de facturation pour les médecins des provinces et des territoires dans lesquels sont recueillies des informations démographiques, la spécialité des intervenants, les raisons de la visite, le diagnostic et les interventions pour la plupart des visites des patients hospitalisés ainsi que pour les patients externes (Quan et al., 2012). De façon complémentaire, les hôpitaux possèdent également des bases de données dans lesquelles sont colligées une foule de données hospitalières (Quan et al., 2012). Troisièmement, la possibilité de coupler des bases de données par un identifiant personnel unique émis par chaque province ou territoire distingue le Canada (Quan et al., 2012). Des pays, tels que l'Australie et le Japon, ne possèdent pas d'identifiant unique, de sorte que pour lier différentes sources de données entre elles des méthodes probabilistes sont utilisées, ce qui engendre des risques d'erreurs (Quan et al., 2012) dans d'éventuels usages secondaires comme les travaux de recherche.

Il est important de noter que les données de santé sont créées par de multiples intervenants (professionnels de la santé, archivistes médicaux, intervenants communautaires, etc.) et qu'elles sont conservées dans des systèmes d'information très variés (Caron, 2021b). Ces facteurs peuvent influencer l'enjeu d'assurance de qualité des données.

Un des premiers enjeux liés au partage est de s'assurer que les usages primaires sont satisfaits et que les données permettent d'optimiser le fonctionnement du système tant dans sa dimension clinique qu'administrative.

Une question centrale : objectifs de la recherche en santé

Le premier usage secondaire reconnu et accepté est l'utilisation des données à des fins de recherche en santé, car celle-ci joue un rôle fondamental dans le système de santé. Du côté clinique, les résultats qu'elle produit ont des effets structurants sur la société. Dans ce contexte, il est essentiel que les démarches scientifiques puissent se baser sur des données probantes de qualité. La recherche en santé est grandement dépendante de l'existence, de l'accès et de la qualité des données (Council of Canadian Academies, 2015). Les données sur la santé sont

utilisées pour répondre à de multiples questions de recherche qui touchent plusieurs aspects de la santé de la population comme les services de santé, les résultats des traitements en matière de santé ou encore la qualité des soins (Quan et al., 2012). Par exemple, « l'utilisation secondaire des données sur la santé est courante dans la recherche épidémiologique et l'examen des dossiers médicaux peut être très utile dans les études de santé publique à grande échelle en raison de la richesse des données pré-collectées disponibles » (traduction libre) (Hill et al., 2013, p. 1). Au Québec, les données détenues par la Régie de l'assurance maladie (RAMQ) sont d'un grand intérêt pour la recherche, car elles contiennent des informations compilées sur une longue période pour l'ensemble de la population québécoise (Blais, 2020 ; cité dans Mercure et Pilon-Larose, 2020). La littérature suggère plusieurs retombées bénéfiques qui justifient l'importance accordée à la recherche en santé et constitue, du coup, la principale raison d'être de cet usage secondaire. Le tableau 1 ci-dessous présente quelques retombées potentielles de la recherche en santé.

Tableau 1.

Quelques retombées potentielles de la recherche en santé

Retombées potentielles de la recherche en santé	Auteurs
Compréhension approfondie des maladies et des traitements	Botsis et al., 2010; Auffray et al., 2016; Wayman et Hunerlach, 2019
Efficacité et sécurité du traitement ou pharmacovigilance	Botsis et al., 2010; Auffray et al., 2016; Wayman et Hunerlach, 2019; Bernelin, 2019; Béranger, 2017; Péchillon, 2015; Girard et Polton, 2018; Gruson et Grass, 2020
Amélioration de la qualité des soins, comme un diagnostic plus rapide	Botsis et al., 2010; Wayman et Hunerlach, 2019; Vogus et al., 2020; Yao et al., 2018
Identification efficace des cibles pour les nouveaux traitements et médicaments	Wayman et Hunerlach, 2019; Bardhan, Chen et Harahanna, 2020

Conception d'essais cliniques pour les populations cibles et réduction du délai de mise en marché des nouveaux traitements	Wayman et Hunerlach, 2019; Bardhan, Chen et Karahanna, 2020
Efficacité opérationnelle et économique de la prestation des soins de santé et planification des effectifs	Botsis et al., 2010; Yang, Ye et Bandarchi, 2019; Bardhan, Chen et Karahanna, 2020; Biedma et Bourret, 2015; Vogus et al., 2020; Hawthorne et Richards, 2017.
Mise en œuvre de la médecine personnalisée - les bons traitements pour les bons patients	Wayman et Hunerlach, 2019; Sarradon-Eck, 2019; Hawthorne et Richards, 2017
Incidences socio-économiques sur la santé, la santé publique, le bien-être et les soins de santé	Botsis et al., 2010

Ces bénéfices potentiels découlent de l'existence de données de qualité et qui puissent être partagées. Pour atteindre cet objectif, les décideurs publics orchestrent des solutions qui répondent aux besoins du fonctionnement et de la performance du système (usages primaires) et les arriment avec les besoins de la recherche dans un cadre d'acceptabilité sociale. Il y a donc de nombreux défis.

Un équilibre à atteindre entre risques et opportunités

À la lecture du tableau 1, on comprend que si les données de santé sont indispensables à la recherche et à la planification et peuvent permettre la réalisation de plusieurs retombées positives ; elles ont la particularité de pouvoir offrir ces bénéfices dans la mesure où les chercheurs sont habilités à les partager, les relier et les réutiliser (Cheung, 2020). Ceci ouvre la porte à de multiples utilisations et à des enjeux de traçabilité dans l'utilisation du renseignement d'un individu en particulier. C'est dans cette perspective que de nombreux auteurs ont abordé le

concept de « contrat social ». Ancré dans un idéal de solidarité, ce dernier est construit sur la base que les individus ont un devoir éthique et moral de partager leurs données de santé pour la recherche afin de soutenir le bien-être collectif (Ballantyne and Schaefer, 2018 ; Prainsack et Buyx, 2017 ; Wellcome Trust, 2015).

Cependant, comme mentionné, l'enjeu réside aussi dans le fait qu'il en découle des risques importants, entre autres, pour la vie privée.

Les données de santé constituent sans aucun doute un terreau fertile tant du point de vue scientifique que médical et économique. Mais les risques associés sont élevés et obligent à veiller à ce que les finalités d'innovation soient compatibles avec les exigences d'une société démocratique, soucieuse de préserver la vie privée et l'autonomie de ses citoyens. Cette ambivalence contraint à faire preuve d'une imagination renouvelée en matière technique et juridique, tant sur le recueil et la conservation que sur les traitements et les usages des données de santé (Peugeot, 2018, p. 30).

À ce jour, il est encore rare de voir des débats ouverts à ce sujet soit une conversation publique honnête sur les bons et les mauvais côtés de ces usages (Malik, 2021). Cette lacune autour de la politisation de l'enjeu laisse place à l'éclosion de croyances et d'opinions qui ne sont pas nécessairement fondées et qui restreignent les possibilités d'avancement.

Partage des renseignements de santé pour des fins de recherche

État de situation

Pour le Québec, la recherche représente un enjeu stratégique et joue un rôle capital pour son développement ainsi que sur son leadership en matière de science et d'innovation (Quirion et Fonds de recherche du Québec, 2020). Au Québec, les banques de données pour la recherche en santé sont principalement détenues par deux instances : la Régie de l'assurance maladie du Québec et le ministère de la Santé et des Services sociaux. Au cours des dernières années, les médias ont rapporté à maintes reprises des irritants quant à l'accès aux données de santé (Mercure et Pilon-Larose, 2020 ; Plamondon Emon, 2017 ; Meloche-Holubowski, 2020 ; Malboeuf, 2021). De plus, du côté de l'usage plus administratif, bien que présent avant, la pandémie de la COVID-19 a montré l'importance d'avoir accès à des données de santé pour éclairer la prise de décisions (Quirion et Fonds de recherche du Québec, 2020). Dans tous les cas, l'accès aux données dans la situation actuelle montre des délais importants (Mercure et Pilon-Larose, 2020 ;

Plamondon Emon, 2017) et une lourdeur administrative (Meloche-Holubowski, 2020 ; Plamondon Emon, 2017) souvent difficile à surmonter. À ce titre, Quirion et les Fonds de recherche du Québec rapportent « (...) des délais d'attente pouvant parfois atteindre 2 ans et plus » (2015, p. 9) pour l'obtention de données de santé. De plus, contrairement aux autres provinces canadiennes, les chercheurs qui souhaitent obtenir des données sans le consentement des usagers (ex. : RAMQ) doivent avoir l'autorisation de la part de deux paliers décisionnels différents, soit la Commission d'accès à l'information ainsi que l'organisme public qui détient les données, et ce, en vertu de la Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels (A-2.1) (Quirion et Fonds de recherche du Québec, 2015).

Ces contraintes constituent un frein à la recherche et à l'innovation et conséquemment à l'amélioration des soins et du système de santé et de services sociaux. Ceci est particulièrement préoccupant alors que les avancées technologiques en santé, dont les développements liés à l'intelligence artificielle, carburent essentiellement aux données. Conscient de ces enjeux, le ministère de la Santé et des Services sociaux a lancé un vaste exercice de consultation visant à revoir la façon dont sont gérées les données de santé. Ce chantier se déroule en amont du dépôt d'un projet de loi prévu pour l'automne 2021 (Lévesque, 2021).

Défis et enjeux

La littérature scientifique fait état de nombreux défis et enjeux concernant le partage des données de santé ce qui fait écho à la situation du Québec. Ceux-ci sont divisés en 5 grandes catégories : l'accès aux données, la qualité des données, les usages des données, la protection du citoyen et la confiance et l'acceptabilité sociale.

L'accès aux données de santé

Obtenir la permission d'accéder aux données de santé semble être un défi dans plusieurs cas. Le processus de demande de données et les délais d'attente pour leur réception varient fortement d'une province à l'autre (Doyle et al., 2020). Certains éléments permettent d'en expliquer la complexité. En effet, les travaux du Council of Canadian Academies (2015) montrent que les plus grands obstacles à la recherche sont institutionnels. Il s'agit notamment de l'application et de l'interprétation différentes et, dans certains cas, « trop prudentes », de la législation sur la protection de la vie privée jumelée à des processus d'approbation complexes et longs qui

entravent l'accès des chercheurs aux données. Dans le même ordre d'idées, Ford et al. (2019) notent l'aversion au risque des décideurs qui sont enclins à interpréter la législation avec prudence par manque de connaissance de ce que la loi autorise. De plus, le manque de gestionnaire de données compétent est aussi un des facteurs qui contribue à la sous-utilisation des données de santé (Pisani et al., 2016).

Qui plus est, les silos institutionnels représentent une barrière supplémentaire à l'accès aux données. Des travaux de recherche ont montré le nombre important de systèmes d'information qui coexistent au sein du réseau de la santé et des services sociaux (Limoges, 2020) ainsi que les formats variés de compilation et de conservation des données des usagers (Caron, 2021a). Ces nombreuses bases de données non uniformisées et souvent non interopérables rendent ardu l'accès aux données de santé par les chercheurs. Les institutions ne sont, toutefois, pas les seules à fonctionner en vase clos. Les chercheurs aussi travaillent en silo ce qui réduit les chances de maximiser la portée des données de santé (Ford et al., 2019). Une plus grande diffusion de la recherche financée par les fonds publics permettrait de mieux octroyer le financement à disposition et assurerait une reddition de compte plus complète aux citoyens quant à l'usage secondaire fait de leurs données.

Les caractéristiques des données de santé

L'utilisation (et la réutilisation) efficace de données de santé peut parfois être difficile en raison de la qualité de celles-ci. En effet, la qualité et la structure des données de santé disponibles sont souvent inégales ce qui représentent un défi majeur pour la recherche (Auffray et al., 2016). Des données incomplètes ou inexactes provoquent des résultats de recherche erronés et inutilisables (Hawthorne et Richards, 2017). De plus, n'étant pas créées d'emblée pour servir la recherche, les données de santé collectées dans un objectif de prestation de soins sont rarement prêtes pour une utilisation secondaire (Ancker et al., 2011; Ford et al., 2019). Elles sont souvent de qualité variable et il arrive fréquemment qu'elles soient incomplètes (Ancker et al., 2011). Un traitement manuel peut alors être nécessaire pour évaluer la qualité et normaliser les formats de données à des fins d'analyse (Ancker et al., 2011). Il existe d'ailleurs un grand nombre d'ouvrages consacrés à l'évaluation de la validité des données de santé (Benchimol et al., 2015).

Par ailleurs, la mise en commun de données provenant de sources différentes engendre des défis au niveau méthodologique. En ce sens, la standardisation des données est un élément central de la qualité de celles-ci tant pour l'utilisation primaire que secondaire. Afin de pouvoir les comparer entre elles et en tirer des conclusions, celles-ci doivent être créées selon des normes communes. Pour y arriver, il est nécessaire de concevoir et d'appliquer des normes et des protocoles uniformes (Council of Canadian Academies, 2015). La gouvernance des données a, ici, un rôle fondamental à jouer pour mettre en place des mécanismes rigoureux à cet égard.

D'ailleurs, la recherche montre que les bénéfices de l'utilisation des données de santé sont décuplés lorsque les données peuvent être liées à de multiples sources (Ford et al., 2019). La combinaison et la mise en commun de données provenant de plusieurs bases sont un avantage considérable pour la recherche ce qui est maintenant plus facile à accomplir avec les capacités de calcul et d'analyse accrues par l'usage de technologies numériques (Lowrance, 2003). Cela dit, les données liées renferment un plus haut risque d'identification des personnes, et, par le fait même, d'abus en lien avec l'utilisation de celles-ci (Lowrance, 2003). Ainsi, il apparaît qu'un cadre juridique strict doit être développé avec tous les éléments propres à bien encadrer les pratiques, mais aussi visant à sanctionner les manquements. De plus, certains auteurs suggèrent que les décisions relatives au couplage de données à des fins de recherche doivent équilibrer les droits individuels et les avantages pour la société (Ford et al., 2019). Dans chaque cas, il est nécessaire d'évaluer les risques et les avantages du projet de recherche, la perception qu'en a la société ainsi que l'infrastructure de sécurité (Ford et al., 2019).

Enfin, à l'ère des données de masse, le volume de données à traiter peut parfois être un enjeu pour les chercheurs souhaitant analyser des données de santé. Bien que les outils technologiques soient de plus en plus performants pour effectuer ce genre de tâche, les chercheurs n'ayant pas les connaissances ou les outils nécessaires pour traiter ce fort volume de données risquent d'être dépassés, notamment parce qu'il manque souvent une documentation adéquate expliquant comment utiliser les données disponibles à des fins de recherche (Quan et al., 2012). Pour éviter ce phénomène, les chercheurs et les organismes de santé doivent travailler de concert pour favoriser une utilisation optimale ces données (Quan et al., 2012).

La protection du citoyen

La littérature soulève l'importance de la protection du citoyen à travers les processus de partage des données de santé à des fins de recherche. Les enjeux sont particulièrement importants pour des données de santé, car un traitement inapproprié de ces informations peut entraîner des préjudices aux individus tels que la discrimination en matière d'emploi ou d'assurance, des atteintes à la réputation ainsi que des préjudices psychologiques (Arora, 2019). Une enquête par sondage a montré que les individus éprouvent des inquiétudes quant à des fuites de données potentielles ou à des utilisations malsaines de leurs informations personnelles (Patil et al., 2015). La valeur monétaire élevée des données de santé ainsi que leur caractère hautement personnel (Martin et al., 2017) ont poussé plusieurs juridictions à se prévaloir de législation d'exception pour les données de santé en raison de leur besoin de protection particulier (Arora, 2019). Par ailleurs, la multiplication des bases de données, « (...) leur dispersion entre les mains de nombreuses entreprises plus ou moins habituées à manipuler des données sensibles, induisent mécaniquement une augmentation des risques de fuite de données, par inadvertance – un sous-traitant technique peu exigeant – ou à la suite de manœuvres crapuleuses. Il se passe rarement plus de quelques jours sans qu'on ne découvre une faille de sécurité sur des logiciels clés ou, plus grave, qu'une fuite de données ne soit révélée » (Peugeot, 2018, p. 36).

Les travaux de Scott et al. (2018) ont identifié trois grandes menaces au partage des données de santé sur la vie privée: 1) les risques de perte de données en raison d'un incident ou par malveillance ; 2) la divulgation statistique soit la possibilité d'identifier un individu à travers des recoupements de données; et, 3) les effets potentiellement négatifs de l'utilisation secondaire des données comme la divulgation d'informations sur une personne alors que cette personne aurait préféré ne pas les divulguer. Les principes FAIR (Fair Information Practices Principles) sont utilisés pour guider les décisions en matière de protection de la vie privée, entre autres, par les États-Unis et l'Union européenne (Kaplan, 2014).

L'anonymisation peut aider à atténuer ces menaces. Elle consiste en un « processus qui consiste à éliminer des identifiants tels que les noms et les dates de naissance des données à caractère personnel, rendant ainsi l'identification des sujets hautement improbable » [traduction libre] (Aitken et al., 2016, p. 19). L'anonymat n'est pas un concept binaire. Il s'agit d'un continuum sur lequel une donnée peut se situer (Lowrance, 2006). Bien que l'anonymisation diminue les risques

d'identification, elle n'en demeure pas moins possible (Malik, 2021). Ceci est particulièrement vrai avec la calculabilité dans l'univers numérique. Malgré l'anonymisation, plus les bases de données sont nombreuses, plus les croisements possibles sont grands (Peugeot, 2018). À ce titre, le Commissaire à l'information du Royaume-Uni mentionne qu'en réalité, il est difficile de déterminer si les données sont bel et bien anonymes ou si elles contiennent toujours des informations à caractère personnel (2015; cité dans Aitken et al., 2016, p. 19). Le gouvernement australien a, pour sa part, appris cette leçon en 2016, alors qu'il a publié sous forme de données ouvertes les données de remboursements de dépenses médicales « anonymisées » de 2,9 millions d'Australiens (Peugeot, 2018). À partir de ces mêmes données, la publication scientifique en ligne ScienceX s'est permis de ré identifier rapidement sept célébrités australiennes, dont trois députés et un footballeur (Peugeot, 2018).

Dans ce contexte d'incertitude quant à la possibilité de réidentification, la notion de consentement est importante et elle est largement abordée dans la littérature en lien avec l'utilisation des données à des fins de recherche. Selon Mantovani et Quinn (2014), le consentement est un prérequis au traitement des données en santé. Intrinsèquement lié à des questions éthiques, le consentement implique de fournir aux personnes visées des informations dans un format qu'elles peuvent comprendre leur expliquant les usages prévus pour la recherche et leur permettant de poser des questions afin d'obtenir leur accord (Elger et al., 2010). Les modèles d'éthique en recherche ont illustré que le patient se sent beaucoup plus en confiance envers les institutions et leurs décisions lorsque l'obtention de son consentement est une priorité (Holm, 2015). La recherche montre que les individus ont des perspectives différentes concernant le consentement. Selon les travaux de Hill et al. (2013), certains souhaitent avoir la possibilité de donner leur consentement à chaque utilisation de leurs données. Pour d'autres, un consentement unique qui couvre toutes les recherches futures est acceptable. Enfin, certains sont d'accord à l'idée que des recherches puissent être menées sans consentement explicite. De plus, Hill et al. (2013) rappellent qu'un certain nombre d'études ont suggéré que les individus sont plus disposés à renoncer à leur consentement lorsqu'ils comprennent les enjeux et les retombées de la réalisation de ces études.

L'idée de consentement est largement discutée dans la littérature, mais ne fait pas consensus. Par exemple, les experts ne voient pas tous la notion de consentement du même point de vue. Plusieurs s'entendent sur le consentement comme l'élément crucial pour établir un cadre de

gouvernance sur la recherche en santé, alors que d'autres reconnaissent les faiblesses du consentement puisque les usagers n'ont souvent pas d'agentivité sur les décisions en cours dans la recherche (Kraft et al., 2018). De plus, la notion de bénéfice collectif entre également en jeu. Les données de santé ayant le potentiel d'améliorer le bien-être du plus grand nombre, certains auteurs ont adressé des considérations à cet égard. Par exemple, Ballantyne (2019) stipule que les modèles de consentement individuel sont limités et ne devraient pas éclipser une réflexion plus approfondie sur « (...) l'objectif de la recherche, l'équilibre entre les divers intérêts en jeu (privés, gouvernementaux, communautaires, individuels) et le processus de négociation d'un consensus social sur l'utilisation des données » [traduction libre] (p. 357). De son côté, le Commissaire à l'information du Royaume-Uni (Information Commissioner's Office) mentionne que

l'obtention du consentement pour l'anonymisation des données personnelles peut être logistiquement très onéreuse, par exemple lorsqu'il existe un grand nombre d'enregistrements de dossiers personnels. Elle peut même s'avérer impossible - par exemple lorsque les données personnelles sont anciennes et qu'il n'existe aucun moyen fiable de contacter les personnes concernées. C'est souvent le cas des archives historiques qui peuvent contenir les données personnelles d'individus qui sont encore en vie [traduction libre] (2012, p. 28).

Il faut donc souligner que si le consentement est une pratique intéressante ; elle soulève de nombreux problèmes d'application qui peuvent avoir des conséquences sérieuses sur la qualité des données de recherche.

La confiance et l'acceptabilité sociale

Nombre de recherches se sont intéressées à ce qui rend acceptable pour les usagers le partage des données de santé (Aitken et al., 2016 ; Caron et al., 2020). Plusieurs motifs peuvent pousser des chercheurs à entreprendre une étude dans le domaine de la santé. La confiance des individus fluctue selon l'appartenance des chercheurs. À ce titre, la recherche dans le secteur privé (ex. : pharmaceutique) obtient un degré plus faible d'approbation de la part des individus que des chercheurs universitaires (Caron et al., 2020). De plus, l'attitude favorable face à la recherche augmente s'il y a des bénéfices collectifs (Scott et al., 2018 ; Caron et al., 2020).

La légitimité de la recherche en santé s'appuie grandement sur la confiance des individus envers les organisations et les professionnels, notamment sur la confiance accordée aux professionnels qui ont un devoir de confidentialité (Carter, Laurie et Dixon-Woods, 2015). Les limites de

l'acceptabilité doivent être renégociées en permanence pour s'assurer que les pratiques s'alignent sur les valeurs et les préoccupations des citoyens (Aitken et al., 2016 ; Caron et al., 2020). Ceci est d'autant plus important lorsqu'on considère les relations économiques et commerciales qui sous-tendent à certains types de recherche. Il est à noter que ce qui est juridiquement admissible ne coïncide pas nécessairement avec ce qui est socialement acceptable (Carter, Laurie et Dixon-Wood, 2014).

Enfin, il est à noter que la transparence dont la clarté en ce qui a trait au modèle de partage et aux pratiques est centrale à cette acceptabilité. Le régime de gestion et de gouvernance de l'information joue donc un rôle clé dans l'acceptabilité sociale.

Modèles recensés de partage des renseignements de santé

Les littératures scientifiques et grises suggèrent plusieurs modèles pour favoriser un partage des renseignements de santé avec les chercheurs. Les modèles recensés sont présentés suivis d'exemples concrets de leur utilisation.

Présentation des modèles

Le registre national

Présenté dans les travaux de Grant et al. (2013), le modèle du registre national est utilisé dans certaines juridictions pour gérer l'accès aux données de santé pour la recherche. Dans ce modèle, la technologie permet d'extraire des données de dossiers médicaux électroniques (DMÉ) et de relier ces informations avec un numéro d'identifiant pour chacun des usagers. D'abord, les usagers sont invités à inscrire leurs coordonnées soit via le site web ou un formulaire qui est saisi par le personnel du registre. En inscrivant leurs coordonnées, ils consentent à ce que l'information contenue dans leur DMÉ soit utilisée pour déterminer leur éligibilité à diverses études en fonction des critères recherchés. Ces études ont reçu les approbations requises en matière d'éthique et de gouvernance. Ensuite et de là, les chercheurs communiquent avec les administrateurs du registre en leur indiquant leurs critères de sélection (inclusion et exclusion). Le modèle du registre permet de coupler les données entre les DMÉ et l'identifiant de l'utilisateur afin de créer la banque d'utilisateurs qui répondent aux critères et qui sont disposés à participer à la recherche. Enfin, les usagers dont le dossier répond aux critères sont interpellés par le personnel du registre. S'ils donnent leur approbation, l'équipe de recherche peut communiquer directement avec eux.

Ce modèle est utilisé au Royaume-Uni, plus précisément pour le Scottish Health Research Register (SHARE). Le registre SHARE est encadré par le National Health Service (NHS). Le registre a également une branche SHARE biobanque. Une étude qualitative réalisée à partir de groupes de discussion auprès de plusieurs parties prenantes montre que les patients, les praticiens et les chercheurs sont généralement favorables au concept du registre national pour la recherche (Grant et al., 2013). Certains patients ont, toutefois, exprimé le besoin d'obtenir plus d'informations sur la sécurité de leurs données ainsi que sur l'imputabilité des utilisateurs de celles-ci (Grant et al., 2013). Les patients favorisent une approche de gestion du registre axée sur les bénéfices collectifs ce qui a conduit les participants à rejeter, à la majorité, la notion d'investissement d'entreprises pharmaceutiques dans l'infrastructure du registre (Grant et al., 2013). Dans leur étude, les auteurs en viennent à la conclusion que pour assurer le succès d'un registre national volontaire, une transparence complète est requise afin de maintenir la confiance des usagers (Grant et al., 2013).

L'Ontario a également mis en place un registre qui s'intitule Registre et Réseau des bons résultats dès la naissance (en anglais BORN). Ce dernier concerne la maternité, les nouveau-nés et les enfants. « BORN recueille, interprète, partage et protège rigoureusement des données de haute qualité essentielles pour faire de l'Ontario l'endroit le plus sûr au monde pour avoir un bébé » (BORN Ontario, s.d.). Le registre soutient la recherche et l'innovation scientifiques en partageant des données de qualité.

Les données ouvertes

Le modèle des données ouvertes s'inscrit dans le concept plus large de la science ouverte ; une philosophie qui promeut une transparence complète dans le partage des données afin d'en faciliter l'accès à au plus grand nombre pour améliorer, entre autres, la recherche scientifique (Rios, Zheng et Zheng, 2020). Le principal facteur qui permet aux données ouvertes sur la santé d'avoir des retombées concrètes découle de la volonté du gouvernement à rendre les données plus accessibles et d'accepter d'agir sur les problèmes révélés par les données (Scrollini, 2018). Les travaux de Martin et Begany (2017), réalisés dans le contexte américain, ont fait ressortir plusieurs bénéfices potentiels des données ouvertes en santé. À ce titre, il est possible de noter : une coordination plus efficace de la santé publique, une amélioration de la qualité, de la rapidité d'accès et de l'utilité des données, un plus grand accès pour les chercheurs, une amélioration de

la littératie en santé et l'atteinte d'un gouvernement plus transparent et juste. À l'opposé, les auteurs notent aussi plusieurs défis à la mise en œuvre de données ouvertes. Ceux-ci se situent au niveau :

- des ressources humaines à affecter à la tâche et des connaissances requises pour l'effectuer;
- de la culture organisationnelle, dont l'organisation en silos;
- des contraintes légales et réglementaires qui entravent la diffusion;
- de la qualité des données qui doit être propice à la diffusion à grande échelle;
- technique et de la capacité des systèmes et plateformes informatiques ;
- de la volonté politique.

De plus, pour certains, l'accès ouvert aux données de santé fait naître des craintes quant à de possibles dérives du système pouvant conduire à une marchandisation de l'information (Péchillon, 2015, p. 646).

La protection de la vie privée et de la confidentialité suscite des préoccupations, surtout en ce qui concerne les bases de données sur les soins de santé et la santé publique. On reconnaît que l'appel à l'ouverture et à la transparence doit être tempéré par la nécessité de respecter la vie privée et la confidentialité (Huston, Edge et Bernier, 2019, p. 280).

Ainsi, bien que les jeux de données soient utiles à la recherche, ils ont aussi certaines limites concernant les recherches auxquels ils peuvent servir. À titre d'exemple, le National Health Service du Royaume-Uni (NHS) est généralement réticent à diffuser des données à des fins secondaires, sauf dans le cadre de processus de gouvernance de l'information hautement contrôlés et réglementés, malgré des priorités en matière de données ouvertes au niveau national (Sexton et al., 2018). Dans le cas des données ouvertes, l'utilisation de données anonymisées dans l'intérêt public réduit l'obligation d'obtenir le consentement (Sexton et al., 2018).

Malgré ces défis, plusieurs gouvernements ont développé des plateformes de données dédiées à la santé ou ayant une section portant sur la santé. À titre d'exemple, le gouvernement de l'Ontario par son service de santé publique diffuse des jeux de données sur son portail dédié (Santé publique Ontario, 2021). Au niveau canadien, le Canada diffuse un certain nombre de bases de données gérées par l'Agence de la santé publique, dont sur l'Infobase de données sur la santé

publique (Huston, Edge et Bernier, 2019). Le gouvernement américain procède également à la diffusion de jeux de données sur son portail (Health Data Gov., s.d.). Sous la direction du Department of Health and Human Services, ce portail de données ouvertes vise à rendre accessibles les données aux entrepreneurs, aux chercheurs et décideurs afin d'améliorer les soins de santé (Health Data Gov., s.d.).

Le modèle fiduciaire ou le tiers de confiance

Dans un rapport de 2017 intitulé *Growing the Artificial Intelligence Industry in the UK*, le gouvernement britannique a présenté le concept de fiducie des données comme étant des institutions susceptibles de soutenir le partage des données tout en veillant à l'intérêt public (Hall et Pesenti, 2017). Les fiducies de données agissent comme gestionnaires indépendants des données en valorisant un partage juste, sécurisé et équitable de cette ressource (Milne, Sorbie et Dixon-Woods, 2021). Elles ont pour objectif de prendre des décisions éclairées sur l'accès aux données et les conditions de cet accès en fonction des retombées possibles (Milne, Sorbie et Dixon-Woods, 2021). Ce modèle est ancré dans une gouvernance participative misant sur des processus décisionnels délibératifs. Conceptuellement, la structure juridique et de gouvernance sous-jacente au modèle fiduciaire doit obliger les fiduciaires à représenter et à privilégier les droits et les besoins des fournisseurs de données (bénéficiaires) dans le cadre de la négociation de contrats d'accès et d'utilisation de leurs données auprès, entre autres, d'entreprises privées (Zarkadakis, 2020). Les fiduciaires se voient attribuer un pouvoir discrétionnaire sur les intérêts du bénéficiaire (Arora, 2019). En général, dans une relation fiduciaire, l'agent fiduciaire agit "pour le compte" du bénéficiaire (Arora, 2019). La législation entourant la mise en place de ce modèle devrait aussi s'assurer de contrecarrer la relation de pouvoir asymétrique entre les fiduciaires et les bénéficiaires et protéger le bénéficiaire contre l'opportunisme (Licht, 2016).

Pour le domaine de la recherche en santé, une fiducie de données peut servir d'organisme de coordination qui assure la gestion des demandes d'accès pour des données destinées à la recherche visant, par exemple, l'amélioration de la prestation de soins (Element AI et Nesta, 2019). Certains écrits suggèrent la mise en place de multiples fiducies, chacune étant représentative des besoins d'un groupe particulier de parties prenantes en raison des besoins particuliers et contextuels inhérents au domaine de la santé (Element AI et Nesta, 2019). Ce modèle offre une grande flexibilité et favorise une gestion éclairée selon les thématiques.

Parmi les cas d'utilisation de ce modèle dans le domaine de la santé, notons le Cincinnati Children's Hospital Medical Center qui utilise le modèle fiduciaire pour assurer une collaboration balisée et efficace entre un réseau de recherche pédiatrique et le réseau de soins. Cela permet à cette instance hospitalière la mise en commun des données et un progrès des recherches médicales (Gouvernement du Canada, 2020). De plus, à l'Institut médical Johns Hopkins, le modèle de fiduciaire fournit l'infrastructure technique, les normes, les politiques et procédures nécessaires pour rassembler les données relatives aux patients et aux membres de tout le système de santé afin de soutenir la réalisation de leur mission organisationnelle (Johns Hopkins Institute for Clinical and Translational Research, s.d.).

Le National Health Service de l'Écosse (NHS Scotland) a aussi des plateformes nommées Data Safe Haven qui fournissent un accès aux données de santé pour la recherche tout en protégeant la confidentialité de celles-ci. Les Data Safe Haven ont été créés pour répondre à la nécessité d'avoir un accès rapide à des données santé de haute qualité à des fins de recherche (NHS Scotland, 2021). Il s'agit d'un environnement sécurisé, soutenu par un personnel formé et des processus balisés, dans lequel les données de santé peuvent être traitées et reliées à d'autres données relatives à la santé ou non. Les données sont mises à disposition sous une forme dépersonnalisée afin d'être analysées (NHS Scotland, 2021). Les données demeurent sous le contrôle du NHS et sont conformes à la législation ainsi qu'aux politiques du NHS (NHS Scotland, 2021).

Les banques de données liées (en anglais linked databank)

Les systèmes d'information captent une foule de données au sujet des usagers dans des dossiers médicaux électroniques épars (Caron, 2021a). Les systèmes de santé sont souvent composés de multiples intervenants stockant les données des usagers dans des divers systèmes d'information non interopérables (Lyons et al., 2009). Pour faire face à cette problématique et ainsi valoriser les données recueillies à des fins de recherche, certaines juridictions ont décidé de créer des banques de données liées. Ces instances rassemblent des données anonymisées qui peuvent être exploitées à des fins de recherche.

À titre d'exemple, le Western Australia Department of Health a mis sur pied une entité appelée Western Australia Data Linkage. Cette instance vise à créer et maintenir des liens au sein d'un large éventail d'ensembles de données. On dit que ces liens sont créés dans le respect de rigoureux protocoles internationalement reconnus de préservation de la vie privée (Data Linkage

Western Australia, 2021a). Les données et leurs couplages sont soumis à des mécanismes d'appariement probabiliste, à des examens administratifs et de contrôle de la qualité qui a pour finalité la création d'un système robuste et complexe liant des informations individuelles (Data Linkage Western Australia, 2021a). Les données liées peuvent être demandées pour réaliser des projets de recherche, de planification et d'évaluation approuvés sur le plan éthique et présentant des avantages pour la collectivité australienne (Data Linkage Western Australia, 2021a). Un processus d'approbation multipartites assez complexe est en place pour assurer le traitement des demandes d'accès (Data Linkage Western Australia, 2021b).

L'Australie n'est pas le seul endroit qui a mis en place un tel modèle de partage des données de santé. Le Royaume-Uni, par le biais du Secure Anonymised Information Linkage (SAIL) Databank possède une banque de données sécurisée et anonymisée contenant des données servant à la recherche afin d'améliorer la santé et les services (SAIL Databank, 2021). Soutenue et approuvée par le gouvernement, la banque de données SAIL reçoit un financement de la part du gouvernement de la région de Wales (SAIL Databank, 2021). La banque de données SAIL contient des milliards de dossiers individuels pouvant être reliés entre eux pour répondre à d'importantes questions de recherche (SAIL Databank, 2021). Cette plateforme de couplage de données est une importante ressource pour les données sur la population, la santé et les soins sociaux du Royaume-Uni destinées uniquement à la recherche (SAIL Databank, 2021). Elle met en place une structure de protection de la vie privée de type *privacy-by-design* et sa gouvernance est certifiée selon les standards ISO. Une étude visant à identifier des leviers pour bonifier le modèle du SAIL Databank émet plusieurs recommandations à considérer (Mourby et al., 2019). Tout d'abord, il importe que la législation soit claire quant au couplage des données administratives de santé pour la recherche visant l'intérêt public. Ceci dans le but de réduire les négociations et le temps d'attente nécessaire à l'obtention de celles-ci. En deuxième lieu, la recherche suggère que les fournisseurs de données devraient adopter des approches fondées sur les principes *privacy by design* dès la création des données. Troisièmement, les citoyens devraient être mobilisés en continu pour être à l'affût des conditions qui rendent acceptable le partage de leurs données et afin de promouvoir l'importance de la recherche dans le secteur de la santé.

La France a aussi un système national des données de santé (SNDS). Le SNDS a pour finalité de pouvoir nourrir la recherche en santé qui sert l'intérêt public. Ce dernier est sous la gestion de la Caisse Nationale de l'Assurance Maladie des Travailleurs Salariés (CNAMTS). Le SNDS permet de

lier les données de l'Assurance Maladie, les données des hôpitaux, les causes médicales de décès, les données relatives au handicap et un échantillon de données en provenance des organismes d'Assurance Maladie complémentaire (SNDS, 2017).

Les partenariats

Public-privé (PPP)

L'arrivée des données massives, certaines juridictions se sont tournées vers des partenariats avec le secteur privé pour saisir les opportunités offertes par les données de santé. Les PPP sont des accords de coopération entre deux ou plusieurs intervenants des secteurs public et privé (Ballantyne et Stewart, 2019). À titre d'exemple, l'Union européenne a lancé un projet visant à créer des « espaces d'innovation » (i-Spaces) (European Commission, 2014). Ces espaces offrent un environnement sécurisé propice à la collaboration et l'innovation intersectorielles à l'aide de données commerciales et publiques (European Commission, 2014). Un autre exemple est la collaboration entre le NHS au Royaume-Uni et Google DeepMind AI. À ce titre, le NHS a fourni des données de santé à Google afin que l'entreprise puisse effectuer la recherche nécessaire au développement d'une application de diagnostic pour les maladies rénales (Ballantyne et Stewart, 2019). Comme le rapportent la BBC (2017) et les travaux de Powles et Hodson (2017), ce partenariat ne s'est pas fait sans remous. Les enjeux légaux en ce qui a trait la protection de la vie privée demeurent au cœur du débat.

Comme le démontre cet exemple, les partenariats avec le secteur privé sont particulièrement délicats. Ils soulèvent des questionnements importants du côté de l'acceptabilité sociale, de la marchandisation des données et de la propriété de celles-ci. Les travaux de Ballantyne et Stewart (2019) suggèrent la mise en place d'un régulateur (en anglais *stewart*) responsable de veiller à promouvoir une utilisation appropriée des données, tout en faisant valoir les intérêts des personnes concernées ainsi que d'un cadre de gouvernance rigoureux.

Universitaire

Dans certains contextes, il peut s'avérer judicieux de faire un partenariat avec des établissements d'enseignement. C'est ce type de modèle que le ministère de la Santé du Manitoba a choisi. En effet, à travers une entente contractuelle avec l'Université du Manitoba, le Manitoba Centre for Health Policy (MCHP) a été créé (Katz et al., 2020). Les auteurs nous font part des principales

caractéristiques de modèle. Premièrement, les thèmes des projets sont soumis par les chercheurs du MCHP, par les ministères et par les responsables des régions sanitaires provinciales. Deuxièmement, les données proviennent principalement des ministères du gouvernement du Manitoba, notamment le ministère de la Santé, des Familles, de l'Éducation et de la Justice ainsi que de plusieurs autres organismes gouvernementaux interministériels. Troisièmement, le MCHP agit en tant que gardien des données recueillies en continu lors de la prestation des services publics au Manitoba. Quatrièmement, d'autres renseignements peuvent provenir de laboratoires provinciaux, de programmes cliniques, d'organismes de sensibilisation communautaires et sociaux et d'organismes de gouvernance autochtones. Au total, ce sont plus de 90 bases de données qui sont conservées dans le dépôt centralisé de données de recherche sur la population du Manitoba au MCHP.

La recherche de Katz et al. (2020) met en lumière certaines leçons apprises par rapport au fonctionnement du MCHP. Tout d'abord, la confiance des parties prenantes est une condition essentielle de sa réussite. En ce sens, il importe de mettre en œuvre les conditions favorables pour s'assurer de la maintenir. Deuxièmement, la recherche fait ressortir l'importance de travailler étroitement avec les chercheurs qui utilisent les données du début à la fin de la recherche afin de pouvoir s'assurer que les données sont utilisées et interprétées correctement. Enfin, la recherche montre qu'il est primordial de garder des traces des méthodologies utilisées afin de veiller à la cohérence de l'interprétation des données à travers le temps. Des outils d'encadrement ont été créés en ce sens.

Considérations pour un modèle de partage des données de santé avec les chercheurs

En plus des modèles présentés, la littérature identifie des principes et des valeurs à respecter dans le cadre d'un partage des données de santé à des fins de recherche. Les travaux du Council of Canadian Academies (2015), à travers une revue des meilleures pratiques quant au partage des données à des fins de recherche, mettent de l'avant quatre principes à respecter pour assurer le succès de ce type d'initiatives. Premièrement, les instances responsables du partage doivent permettre une utilisation appropriée des données pour améliorer le bien-être public. Le bien collectif et l'intérêt public sont d'ailleurs mentionnés dans plusieurs des modèles recensés. Deuxièmement, tout en reconnaissant que les risques ne peuvent être entièrement éliminés, ceux-ci doivent être gérés en commençant par identifier l'éventail de risques possibles, en

encadrant l'accès aux données et en minimisant ces risques dans la mesure du possible. Troisièmement, il est important de veiller au respect de la vie privée des citoyens en réduisant au maximum les risques et en rassurant les citoyens sur la manière dont ils sont protégés. Enfin et quatrièmement, le maintien de la confiance publique est primordial. Pour ce faire, il faut démontrer la fiabilité, notamment en utilisant les données de manière appropriée et en démontrant la valeur sociale de la recherche qui en résulte.

De façon complémentaire, les travaux de Ballantyne (2019) proposent d'appliquer les valeurs fondamentales en santé publique à la recherche en santé. Les valeurs mises de l'avant sont les suivantes :

- Le bénéfice public soit l'intégrité scientifique et la valeur sociale produite ;
- La proportionnalité soit la nécessité de la recherche et la réduction des risques associés ;
- L'équité ;
- La confiance ;
- La redevabilité soit la transparence et la reddition de comptes.

Discussion et conclusion²

Il faut d'abord rappeler que les modèles recensés existent dans des cadres institutionnels nationaux singuliers. Ceci signifie qu'ils s'insèrent dans un amalgame de lois et règlements propres à chaque cadre institutionnel. Cette variable est essentielle pour comprendre qu'une exploitation de l'un ou l'autre de ces cadres dans un contexte différent pourrait présenter des lacunes ou simplement ne pas fonctionner efficacement pour une multitude de raisons légales, conceptuelles ou encore de politiques publiques et administratives. Ensuite, les modèles recensés présentent des caractéristiques communes et singulières. Les caractéristiques que nous retrouvons sont liées à la protection des renseignements personnels, l'anonymisation, la sécurité ou encore les enjeux de consentement et de cadre de gouvernance. Certains modèles sont, par exemple, plus ouverts et accessibles alors que d'autres sont chapeautés par une gouvernance plus stricte comme le modèle faisant appel au tiers de confiance. Ces différences entraînent des répercussions sur le fonctionnement et ce que peut permettre le modèle. Par exemple, lorsqu'il y a chapeutage par

² Cette section s'appuie entre autres sur les travaux de March (1999), Cyert et March (1992) et Brunsson et Olsen (1993).

une gouvernance bien définie, on observe que la raison d'être de l'accès semble plus claire et les mécanismes en soutien mieux balisés.

L'idéal-type d'un modèle de partage des renseignements de santé en vue d'améliorer la performance d'un système de santé comme celui du Québec repose précisément sur un certain nombre de choix qui reflèteront le modèle institutionnel dans lequel nous sommes. Ici, ce modèle est complexe vu le nombre de ses institutions/organisations, ses multiples acteurs (professionnels, administrateurs, etc.), ses objectifs fragmentés, ses extrants différenciés, les relations entre les institutions et les acteurs ou encore ses divers liens communicationnels technologiques, hiérarchiques ou encore de contrôle. Cette densité et cette intensité rendent les relations complexes et la production de données et d'information fortement décentralisée ce qui peut poser des défis de cohérence. Par exemple, un acteur privé n'a pas nécessairement les mêmes objectifs que ceux de l'acteur public avec qui ou pour qui il travaille. Les défis de partage à des fins autres que pour l'usage premier pour lequel la donnée a été créée sont donc nombreux et sont à la fois techniques, éthiques et conceptuels. Pour y parvenir, il faut un certain degré d'intégration. Cette intégration doit d'abord se faire au niveau de la prestation de soins – usages primaires des données – et ensuite en y ajoutant les usages secondaires. L'intégration ne signifie pas la centralisation, mais plutôt l'assurance d'une cohérence d'ensemble, d'une connectivité bien structurée et suffisante et d'une interdépendance entre les éléments qui composent le système de manière à appuyer la concrétisation de cette cohérence.

Le premier élément à considérer est donc la cohérence qui est essentielle afin de bien orienter les mécanismes de fonctionnement du partage des données de manière qu'ils répondent aux divers enjeux mentionnés/souhaités dans le modèle. Ces éléments incluent, entre autres, les usages permis, les accès, la sécurité, la protection des renseignements ou le consentement au partage des renseignements de santé. À partir de la modulation des multiples instruments légaux et de politiques ainsi que de l'implantation de normes et de pratiques encadrant la création et l'usage des données, la cohérence obtenue sert à orienter le fonctionnement pour les besoins liés aux usages primaires et secondaires. Ici se posent les questions de mission, d'objectifs, d'extrants à produire et de résultats attendus pour le modèle de partage. Pour permettre la cohérence, il faut que les objectifs liés au partage (usages secondaires) soient conséquents avec ceux liés aux usages primaires. Autrement dit, on ne crée pas de données spécifiquement pour des usages

secondaires, mais ces derniers se nourrissent des données créées pour la prestation de soins et servent cette dernière. Pour assurer la cohérence, il faut qu'une telle loi soit construite à partir des objectifs de la Loi sur les services de santé et les services sociaux. Une loi sur le partage des renseignements de santé doit donc être ancrée directement à l'intérieur des objectifs de la Loi sur les services de santé et les services sociaux et être énoncée en termes de sous-objectifs à ceux de cette dernière. De cette manière, les mécanismes d'application seront intégrés et complémentaires à ceux qui assurent le fonctionnement du système de santé et de services sociaux.

Ceci nous amène au second élément soit l'interdépendance. Actuellement, il y a beaucoup d'indépendance entre les éléments du système, et ce, à plusieurs égards. Par exemple, dans les trajectoires de soins (Caron, 2021a ; 2021b) on trouve une multitude de professionnels et de sous-systèmes qui opèrent avec des objectifs singuliers comme la production d'un diagnostic, d'une image médicale, d'une chirurgie, etc. Le tout sert l'objectif plus global de la prestation de soins, mais plusieurs parties fonctionnent de manière indépendante et autoportante. Bien que ces activités soient liées les unes aux autres, lorsqu'on les analyse du point de vue de la prestation de services aux usagers et de leur santé, elles demeurent mal sinon peu intégrées au sens de March (1999). L'application du critère d'interdépendance permet d'améliorer voire de garantir l'existence de séquences d'activités multidirectionnelles et qui nécessitent (forcent) l'interopérabilité par les liens de causalité réciproque qui les caractérisent.

Enfin, le troisième élément est celui de la connectivité structurelle. Ici, on comprend l'importance de connaître la nature des relations entre les différentes parties du système de santé de manière à s'assurer qu'elles communiquent de manière appropriée entre elles. Quelle est la nature de la relation entre le diagnostic clinique de l'omnipraticien et la prescription fournie par le pharmacien ? Quelle est la nature de la relation entre cette prestation de médicament et la physiothérapie ? Pour que les données destinées aux usages secondaires en recherche en santé soient optimales, elles doivent refléter leurs contextes de signification de la manière la plus précise et exhaustive possible. La donnée pharmaceutique est plus intéressante si elle informée par la donnée liée au diagnostic médical et l'état de santé général du patient. De même, la pratique réflexive du clinicien s'enrichira de la donnée sur le médicament. Pour la recherche, un modèle bien intégré des flux d'information au niveau des usages primaires aura déjà permis

d'établir les liens cruciaux non pas sur la base de modèles théoriques ou hypothétiques, mais sur la base de la pratique et de la nature concrète des relations entre les différentes parties du système. Pour la pratique et la performance du système, la connectivité structurelle appuie directement la prestation directe de soins. Son élargissement à la recherche permet en plus d'inclure de manière mesurable la contribution de la recherche à la performance du système en en faisant une variable aux liens de causalité réciproque avec les usages primaires c'est-à-dire qui découlent de l'exercice des soins de santé et services sociaux. De la sorte, la recherche devient une partie du modèle de performance et nourrit le système de santé en vue de l'améliorer et de le rendre plus performant. D'ailleurs, en en faisant un élément qui est partie intégrante du système, on légitimise la recherche et on accroît l'acceptabilité sociale par une reddition de comptes plus limpide.

On comprend que les principes qui guideront le partage des renseignements de santé ne peuvent être énoncés qu'en pure symbiose avec les objectifs de la Loi sur les services de santé et les services sociaux. Le modèle de partage qui sera choisi par la suite n'en sera que plus robuste et efficace. Pour être viable et servir la performance du système de santé (clinique et administrative), le modèle de partage des renseignements de santé devrait établir les principes directeurs qui assurent l'intégration en fonction des grands objectifs de la Loi sur les services de santé et les services sociaux

Références

- Aitken, M., de, S. J. J., Pagliari, C., Jepson, R. et Cunningham-Burley, S. (2016). Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *Bmc Medical Ethics*, 17(1), 1-24.
- Ancker, J. S., Shih, S., Singh, M. P., Snyder, A., Edwards, A., et Kaushal, R. (2011). Root causes underlying challenges to secondary use of data. *Amia Annual Symposium Proceedings, 2011*, 57–62.
- Arora, C. (2019). Digital health fiduciaries: protecting user privacy when sharing health data. *Ethics and Information Technology*, 21(3), 181–196.
- Auffray, C., Balling, R., Barroso, I., Bencze, L., Benson, M., Bergeron, J., Bernal-Delgado, E., Blomberg, N., Bock, C., Conesa, A., Del, S. S., Delogne, C., Devilee, P., Di, M. A., Eijkemans, M., Flicek, P., Graf, N., Grimm, V., Guchelaar, H. J., ... Zanetti, G. (2016). Making sense of big data in health research: towards an eu action plan. *Genome Medicine*, 8(1), 1-13.
- Ballantyne, A. et Schaefer, G.O. (2018) Consent and the ethical duty to participate in health data research. *Journal of Medical Ethics*, 44, 392–396.
- Ballantyne, A. (2019). Adjusting the focus: a public health ethics approach to data research. *Bioethics*, 33(3), 357–366.
- Ballantyne, A. et Stewart, C. (2019). Big data and public-private partnerships in healthcare and research : the application of an ethics framework for big data in health and research. *Asian Bioethics Review*, 11(3), 315–326.
- Bardhan, I., Chen, H. et Karahanna, E. (2020) Connecting systems, data, and people: A multidisciplinary research roadmap for chronic disease management, *MIS Quarterly*, 44(1), 185-200.
- BBC (2017, 3 juillet). Google DeepMind NHS app test broke UK privacy law. BBC News. Repéré à <https://www.bbc.com/news/technology-40483202>
- Benchimol, E. I., Guttman, A., Smeeth, L., Harron, K., Langan, S. M., Moher, D., Peteresen, I., Sorensen, H. T., et von, E. E. (2015). The reporting of studies conducted using observational routinely-collected health data (record) statement. *Plos Medicine*, 12(10), 1-22.
- Béranger, J. (2017). Chapitre 1. De l'étude des risques à la traduction des enjeux éthiques des Big Data en santé. *Journal international de bioéthique et d'éthique des sciences*, 3(28), 15-25.
- Bernelin, M. (2019) Intelligence artificielle en santé: la ruée vers les données personnelles, *Cités*, 4(80), 75-89.
- Biedma, J. M. et Bourret, C. (2015) Les enjeux des projets d'interopérabilité en santé pour la mobilité des citoyens. Le cas de l'Espagne, *Projectics / Proyéctica / Projectique*, 1(13), 23-36.

- BORN Ontario (s.d.). *About BORN*. Repéré à <https://www.bornontario.ca/en/about-born/about-born.aspx>
- Botsis, T., Hartvigsen, G., Chen, F., et Weng, C. (2010). Secondary use of EHR: data quality issues and informatics opportunities. *Summit on Translational Bioinformatics, 2010*, 1–5.
- Brunsson, N., et Olsen, J. P. (1993). *The reforming organization*. Routledge.
- Caron, D. J. avec la collab. de Lamontagne, R., Drouin, O., Nicolini, V. et Bernardi, S. (2021a). *Partage des données en santé au Québec : modélisation et cartographie des trajectoires de partage des données*. Rapport de recherche. Chaire de recherche en exploitation des ressources informationnelles. Gatineau : École nationale d'administration publique.
- Caron, D. J. avec la collab. de Nicolini, V., Bernardi, S. et Lamontagne, R. (2021b). *L'utilisation des systèmes informationnels comme levier à l'amélioration de la performance dans les trajectoires usagers*. Rapport de recherche. Chaire de recherche en exploitation des ressources informationnelles. Gatineau : École nationale d'administration publique.
- Caron, D. J., Montmarquette, C., Prud'homme, A., Bernardi, S. et Nicolini, V. (2020). *Projet sur l'acceptabilité sociale du partage des renseignements de santé - Enquête sur l'acceptabilité sociale du partage des renseignements de santé : constatations, résultats et variations*. Rapport final. Chaire de recherche en exploitation des ressources informationnelles. Gatineau : École nationale d'administration publique.
- Carter P., Laurie G.T., Dixon-Woods, M. (2015). The social licence for research: why *care.data* ran into trouble. *Journal of Medical Ethics, 41*, 404-409.
- Cheung, S. (2020). Disambiguating the benefits and risks from public health data in the digital economy. *Big Data & Society, 7*(1),1-15.
- Council of Canadian Academies (2015). Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation. *Accessing health and health-related data in Canada*. Council of Canadian Academies.
- Cyert, R. et March, J. G. (1992). *A Behavioral Theory of the Firm*, Cambridge: Wiley BlackWell.
- Data Linkage Western Australia (2021a). *About Data Linkage WA*. Repéré à <https://www.datalinkage-wa.org.au/about/>
- Data Linkage Western Australia (2021b). *Data Application Guide*. Repéré à https://www.datalinkage-wa.org.au/wp-content/uploads/2021/02/Data_application_guide.pdf
- Despotou, G., Jill Evans, J. et al. (2020). Evaluation of patient perception towards dynamic health data sharing using blockchain based digital consent with the Dovetail digital consent application: A cross sectional exploratory study, *Digital Health, 6*, 1–11.
- Doyle, C. M., Lix, L. M., Hemmelgarn, B. R., Paterson, J. M., et Renoux, C. (2020). Data variability across canadian administrative health databases: differences in content, coding, and completeness. *Pharmacoepidemiology and Drug Safety, 29*, 68–77.

- Element AI et Nesta (2019). *Data Trusts A new tool for data governance*. Repéré à https://hello.elementai.com/rs/024-OAQ-547/images/Data_Trusts_EN_201914.pdf
- Elger, B. S., Iavindrasana, J., Lo, I. L., Müller H, Roduit, N., Summers, P., et Wright, J. (2010). Strategies for health data exchange for secondary, cross-institutional clinical research. *Computer Methods and Programs in Biomedicine*, 99(3), 230–251.
- European Commission (2014). Public-Private Partnership (PPP) for Big Data. Communiqué de presse. Repéré à http://europa.eu/rapid/press-release_MEMO-14-583_en.html
- Ford, E., Boyd, A., Bowles, J. K. F., Havard, A., Aldridge, R. W., Curcin, V., Greiver, M., Harron, K., Katikireddi, V., Rodgers, S. E., et Sperrin, M. (2019). Our data, our society, our health: a vision for inclusive and transparent health data science in the United Kingdom and beyond. *Learning Health Systems*, 3(3), 1-12.
- Girard, M. et Polton, D. (2018) La nouvelle réglementation sur l'accès aux données de santé et sa mise en œuvre deux ans après la loi de modernisation de notre système de santé, *Journal du Droit de la Santé et de l'Assurance - Maladie (JDSAM)*, 3(20), 25-29.
- Gouvernement du Canada (2020). *Gouvernance du contenu numérique et fiducies de données — Diversité des contenus à l'ère numérique*. Repéré à <https://www.canada.ca/fr/patrimoine-canadien/services/diversite-contenus-ere-numerique/gouvernance-contenu-numerique-fiducies-donnees.html#a5>
- Grant, A., Ure, J., Nicolson, D. J., Hanley, J., Sheikh, A., McKinstry, B., et Sullivan, F. (2013). Acceptability and perceived barriers and facilitators to creating a national research register to enable 'direct to patient' enrolment into research: the Scottish health research register (SHARE). *Bmc Health Services Research*, 13, 1-15.
- Gruson, D. et Grass, E. (2020) Quelle régulation positive éthique de l'intelligence artificielle en santé?, *Les Tribunes de la santé*, 1(63), 25-33.
- Hall, D. W. et Pesenti, J. (2017). *Growing the artificial intelligence industry in the UK*. Rapport indépendant pour le Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy.
- Hawthorne, K. H. et Richards, L. (2017). Personal health records: a new type of electronic medical record. *Records Management Journal*, 27(3), 286–301.
- Health Data Gov. (s.d.). *Health Data.Gov*. Repéré à <https://healthdata.gov/stories/s/jw2e-4hhu>
- Hill, E. M., Turner, E. L., Martin, R. M., et Donovan, J. L. (2013). "let's get the best quality research we can": public awareness and acceptance of consent to use existing data in health research: a systematic review and qualitative study. *Bmc Medical Research Methodology*, 13, 1-10.
- Holm, S. (2015). *Informed Consent: Ethical and Legal Issues*. eLS.
- Huston, P., Edge, V. L., et Bernier, E. (2019). Reaping the benefits of open data in public health. *Canada Communicable Disease Report*, 45(11), 252–256.

- Information Commissioner's Office (2012). *Anonymisation: managing data protection risk code of practice*. Repéré à <https://ico.org.uk/media/1061/anonymisation-code.pdf>
- Johns Hopkins Institute for Clinical and Translational Research (s.d.). Data Trusts. Repéré à https://ictr.johnshopkins.edu/programs_resources/programs-resources/i2c/data-trust/
- Katz, A., Enns, J., Smith, M., Burchill, C., Turner, K., et Towns, D. (2020). Population data centre profile: the Manitoba centre for health policy. *International Journal of Population Data Science*, 4(2), 1-10.
- Kaplan, B. (2014). *How Should Health Data Be Used? Privacy, Secondary Use, and Big Data Sales*. Working Paper No. 14-025 Cambridge: Yale University Institute for Social and Policy Studies.
- Kraft, S., Cho, M., Gillespie, K., Halley, M., Varsava, N., Ormond, K., Luft, H., Wilfond, B., et Lee, S. (2018). Beyond consent: building trusting relationships with diverse populations in precision medicine research. *The American Journal of Bioethics : Ajob*, 18(4), 3–20.
- Lévesque, F. (2021, 8 juin). *Dubé veut faciliter l'accès aux données*. La Presse. Repéré à <https://www.lapresse.ca/actualites/sante/2021-06-08/apres-pandemie/dube-veut-faciliter-l-acces-aux-donnees.php>
- Licht, A. N. (2016). *Motivation, information, negotiation: Why fiduciary accountability cannot be negotiable*. Rochester, NY: Social Science Research Network. Repéré à https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2811237
- Limoges, K. (2020). Le Dossier santé numérique (DSN) : un immense projet de 3 G. *Point en santé et services sociaux*, 15(4), p. 12-14.
- Lipsitz, L. (2012). Understanding health care as a complex system: the foundation for unintended consequences. *JAMA*, 308(3), 243–244.
- Lowrance, W. (2003). Learning from experience: privacy and the secondary use of data in health research. *Journal of Health Services Research and Policy*, 8(3), 2–7.
- Lowrance, W. (2006). *Access to Collections of Data and Materials for Health Research*. Rapport de recherche Medical Research Council (MRC) and the Wellcome Trust. Repéré à <https://wellcomecollection.org/works/sz4ea3nx>
- Lyons, R., Brooks, C., Brown, G., Ford, D., Verplancke, J.-P., John, G., Jones, K., et Leake, K. (2009). The sail databank: linking multiple health and social care datasets. *Bmc Medical Informatics and Decision Making*, 9, 1–8.
- March, J. (1999). A learning perspective on the network dynamics of institutional integration. Dans M. Egeberg et P. Laegreid (Éds) *Organizing Political Institutions* (pp. 129-155). Oslo: Scandinavian University Press.

- Malik, K. (2021, 6 juin). *Tell me how you'll use my medical data. Only then might I sign up*. The Guardian. Repéré à <https://www.theguardian.com/commentisfree/2021/jun/06/tell-me-how-youll-use-my-medical-data-then-i-might-sign-up>
- Malboeuf, M.-C. (2021, 18 juillet). *Données sous haute sécurité*. La Presse. Repéré à <https://www.lapresse.ca/actualites/2021-07-18/acces-aux-donnees-de-recherche/donnees-sous-haute-securite.php>
- Mantovani, E., et Quinn, P. (2014). Mhealth and data protection - the letter and the spirit of consent legal requirements. *International Review of Law, Computers & Technology*, 28(2), 222–236.
- Martin, E. G., et Begany, G. M. (2017). Opening government health data to the public: benefits, challenges, and lessons learned from early innovators. *Journal of the American Medical Informatics Association*, 24(2), 345–351.
- Martin, G., Martin, P., Hankin, C., Darzi, A. et Kinross, J. (2017). Cybersecurity and healthcare: how safe are we? *BMJ*, 358, 1-4.
- Meloche-Holubowski, M. (2020, 11 mai). *Accessibilité aux données sur la santé : « Nous sommes encore à l'âge de pierre »*. Radio-Canada. Repéré à <https://ici.radio-canada.ca/nouvelle/1701245/donnees-sante-publique-pandemie-coronavirus-covid>
- Mercure, P. et Pilon-Larose, H. (2020, 22 août). *Accès aux données de la RAMQ: « injuste », disent des chercheurs*. La Presse. Repéré à <https://www.lapresse.ca/actualites/sante/2020-08-22/acces-aux-donnees-de-la-ramq-injuste-disent-des-chercheurs.php>
- Milne R., Sorbie A. et Dixon-Woods, M. (2021). What can data trusts for health research learn from participatory governance in biobanks? *Journal of Medical Ethics*, 1-6.
- Mourby, M. J., Doidge, J., Jones, K. H., Aidinlis, S., Smith, H., Bell, J., Gilbert, R., Dutey-Magni, P., et Kaye, J. (2019). Health data linkage for uk public interest research: key obstacles and solutions. *International Journal of Population Data Science*, 4(1), 1-13.
- NHS Scotland (2021). *Data Safe Haven*. Repéré à <https://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens>
- Patil, S., Patruni, B., Lu, H., Dunkerley, F., Fox, J., Potoglou, D., et Robinson, N. (2015). *Privacy of health records: Europeans' preferences on electronic health data storage and sharing*. Santa Monica: Rand Corporation.
- Péchillon, É. (2015). L'accès ouvert aux données de santé : la loi peut-elle garantir tous les risques de dérives dans l'utilisation de l'information ? *L'information Psychiatrique*, 91(8), 645–649.
- Peugeot V. (2018). Données de santé : contours d'une controverse. *L'économie Politique*, 80(4), 30–41.
- Plamondon Emon, E. (2017, 6 mai). *Des délais interminables pour accéder aux données*. Le Devoir. Repéré à <https://www.ledevoir.com/societe/science/497882/des-delais-interminables-pour-acceder-aux-donnees>

- Platt, J., Raj, M., et Kardia, S. L. R. (2019). The public's trust and information brokers in health care, public health and research. *Journal of Health Organization and Management*, 33(7-8), 929–948.
- Pisani, E., Aaby, P., Breugelmans, J. G., Carr, D., Groves, T., Helinski, M. et Guerin, P. J. (2016). Beyond open data: realising the health benefits of sharing data. *Bmj (Clinical Research Ed.)*, 355, 1-5.
- Powles, J., et Hodson, H. (2017). Google Deepmind and healthcare in an age of algorithms. *Health and Technology*, 7(4), 351–367.
- Prainsack B. et Buyx A. (2017). *Solidarity in Biomedicine and Beyond*. Cambridge: Cambridge University Press.
- Quan, H., Smith, M., Bartlett-Esquillant, G., Johansen, H., Tu, K., Lix, L., et Hypertension Outcome and Surveillance Team. (2012). Mining administrative health databases to advance medical science: geographical considerations and untapped potential in Canada. *The Canadian Journal of Cardiology*, 28(2), 152–154.
- Quirion et Les fonds de recherche du Québec (2015, septembre). *Orientations gouvernementales pour un gouvernement plus transparent, dans le respect de la vie privée et de la protection des renseignements personnels*. Mémoire. Consultations sur les orientations du gouvernement du Québec.
- Quirion et Les fonds de recherche du Québec (2020, septembre). *Mémoire formulant des commentaires sur le projet de loi 64 – quant à l'accès aux renseignements personnels, à des fins de recherche*. Déposé à la Commission des institutions.
- Rios, R. S., Zheng, K. I. et Zheng, M.-H. (2020) Data sharing during COVID-19 pandemic: what to take away. *Expert Review of Gastroenterology & Hepatology*, 14(12), 1125-1130.
- SAIL Databank (2021). Overview. Repéré à <https://saildatabank.com/about-us/overview/>
- Santé publique Ontario (2021). *Données ouvertes*. Repéré à <https://www.publichealthontario.ca/fr/data-and-analysis/using-data/open-data>
- Sarradon-Eck, A. (2019) Le patient contemporain. *Cancer(s) et psy(s)*, 1(4), 51-60.
- Scott, K. Burall, S., Perrin, N., Shelton, P., White, D., Irvine, G. et Grant, A. (2018). Data for Public Benefit Balancing the risks and benefits of data sharing. Repéré à https://www.involve.org.uk/sites/default/files/field/attachemnt/Data%20for%20Public%20Benefit%20Report_0.pdf
- Scrollini, F. (2018). Open data for transparency and accountability in health service delivery: What's new in the digital age?. Repéré à <https://www.u4.no/publications/open-data-for-transparency-and-accountability-in-health-service-delivery-whats-new-in-the-digital-age>
- Sexton, A., Shepherd, E., Duke-Williams, O., et Eveleigh, A. (2018). The role and nature of consent in government administrative data. *Big Data & Society*, 5(2), 1-17.

- Spruit, M. et Lytras, M. D. (2018). Applied data science in patient-centric healthcare: adaptive analytic systems for empowering physicians and patients. *Telematics and Informatics*, 35(4), 643-653.
- Système national des données de santé (2017). *Qu'est-ce que le SNDS ?*. Repéré à <https://www.snds.gouv.fr/SNDS/Qu-est-ce-que-le-SNDS>
- Vogus, T. J., Gallan, A., Rathert, C., El-Manstrly, D. et Strong, A. (2020) Whose experience is it anyway? Toward a constructive engagement of tensions in patient-centered health care, *Journal of Service Management*, 31(5), 979-1013.
- Wayman, C. et Hunerlach, N. (2019). *Realising the value of health care data: a framework for the future*. EY Report. Repéré à https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/life-sciences/life-sciences-pdfs/ey-value-of-health-care-data-v20-final.pdf
- Wellcome Trust (2015) *Enabling Data Linkage to Maximise the Value of Public Health Research Data*. London: Wellcome Trust.
- Yang, B., Ye, S., Bandarchi, M. (2020). A new model for investigating the factors influencing the innovation in medical services. The mediating role of attitude toward knowledge sharing, *VINE Journal of Information and Knowledge Management Systems*, 50(4), 669-690.
- Yao, N., Xi Zhu, Dow, A., Mishra, V. K., Phillips, A. et Tu, SP. (2018). An exploratory study of networks constructed using access data from an electronic health record, *Journal of Interprofessional Care*, 32(6), 666-673.
- Zarkadakis, G. (2020, 10 novembre). *"Data Trusts" Could Be the Key to Better AI*. Harvard Business Review. Repéré à <https://hbr.org/2020/11/data-trusts-could-be-the-key-to-better-ai>

Annexe 1 Liste des mots-clés

Accès données de santé / Health data access
Commercialisation + données de santé / Health data commercialization
Conditions de partage données santé / conditions to share health data
Consentement données de santé / consent health data
Données de santé publiques / public health data
Données ouvertes en santé / open health data
Données personnelles santé / Personal health data
Données santé administratives / Administrative health data
Dossiers médicaux électroniques / Electronic medical records
Enquête sur le partage de données médicales / inquiries on medical data sharing
Entreprise privée données de santé / Private sector health data
E-santé / e-healthcare
Gouvernance données de santé / Health data governance
Gouvernance intelligence artificielle santé / Artificial intelligence governance + health
Health Information Exchange (HIE)
Healthcare data sharing + doctors / partage de données soins de santé + médecins

Identité numérique / Identifiant numérique / Digital identity
Imputabilité données de santé / Accountability health data
Information santé numérique / Digital health information
Modèle de partage de données de santé / Health data sharing model / Health data sharing framework
Partage de données santé / Health data sharing
Perceptions du partage de données professionnels /perceptions data sharing professionals
Perceptions du partage de données patients / perceptions data sharing patients
Personal health record / Dossier santé personnel
Recherche données de santé / Research health data
Santé intelligence artificielle / healthcare artificial intelligence
Santé intelligente / Smart healthcare
Santé numérique / digital health
Transparence partage de données santé / Transparency health data sharing
Volonté de partager données santé /willingness to share health data



Chaire de recherche en exploitation
des ressources informationnelles
