



PARTAGE DE DONNÉES DANS LA RECHERCHE EN ÉDUCATION MÉDICALE AU CANADA

Recommandations de consensus et principes de
gouvernance

JANVIER 2022

TABLE DES MATIÈRES

À propos de ce document	3
Financement	3
Auteurs correspondants	3
Citation	3
Résumé	4
<i>Principes fondamentaux</i>	4
<i>Recommandations</i>	4
Un appel à l'action	5
L'initiative de renforcement des capacités de gestion des données de recherche	6
Notre approche	7
Objectifs et perspectives partagés	9
Avantages de la recherche inter-institutionnelle sur l'éducation basée sur les données	10
Risques de la recherche inter-institutionnelle sur l'éducation basée sur les données	11
<i>Facteurs de risque</i>	11
<i>Risques liés à la procédure</i>	11
Valeurs pertinentes	13
Recommandations	14
<i>Recommandations</i>	14
<i>Principes de gouvernance</i>	15
Annexes	18
<i>A: Consultations et thèmes de la série de pré-ateliers</i>	19
<i>B: Liste des participants</i>	21
<i>C: Aperçu de la série de Connexion de données éducatives</i>	22
<i>D: Concernant la collecte et le partage des données d'identité sociale</i>	23

À PROPOS DE CE DOCUMENT

Les risques, les avantages, les valeurs, les principes et les recommandations pour la recherche éthique, inter-institutionnelle et fondée sur les données en éducation médicale au Canada présentés ici ont été élaborés en collaboration avec des intervenants canadiens en éducation médicale et des gestionnaires de données dans le cadre de 5 ateliers de synthèse des connaissances visant à établir un consensus. Les ateliers ont été complétés par des consultations avant et après l'événement; présentations sur invitation d'experts en partage de données dans les domaines de l'éthique, de l'enseignement supérieur, de la bibliothéconomie et des services d'épidémiologie/santé; et un groupe d'observateurs de représentants d'organisations de défense des médecins stagiaires qui ont veillé à ce que les perspectives des stagiaires soient intégrées dans la déclaration de consensus finale.

FINANCEMENT

Ce travail a été financé par la subvention Connexion du Conseil de recherches en sciences humaines, la bourse de chercheur en éducation du Département de médecine familiale et communautaire (Université de Toronto) et du Département de médecine familiale (Université McMaster).

AUTEURS CORRESPONDANTS

Lawrence Grierson, Ph.D.

Professeur agrégé, Département de médecine familiale

Université McMaster

100, rue Main Ouest, Hamilton, ON, L8P 1H6

(905) 525-9140 x22738

griersle@mcmaster.ca

Kulamakan Kulasegaram, Ph.D.

Professeur agrégé, Département de médecine communautaire et familiale

Président de Temerty, bilan et évaluation des participants

Faculté de médecine de Temerty,

Université de Toronto

200, rue Elizabeth, 1ES-565., Toronto, Ontario, M5G 2C4

(416) 340-3646

mahan.kulaseagram@utoronto.ca

CITATION

Grierson, L., Kulasegaram, K., Button, B., Lee-Krueger, R., McNeill, K., Youssef, A., Cavanagh, A. Partage de données dans la recherche en éducation médicale au Canada : Recommandations de consensus et principes de gouvernance. Hamilton, Ontario : 6 Janvier 2022. Disponible sur : www.dataconnection.ca

RÉSUMÉ

Dans le cadre d'un projet de Connexion parrainé par le CRSH, intitulé : « Initiative de renforcement des capacités de gestion des données de recherche : Éducation des professions de la santé », les intervenants et les gestionnaires de données de l'ensemble de la communauté canadienne de l'éducation médicale ont participé à une série de 5 ateliers successifs de synthèse des connaissances sur invitation seulement qui visaient à faire avancer la réflexion collective et à dégager un consensus en ce qui concerne la collecte et l'utilisation de données sur l'éducation des professions de la santé aux fins de la formation médicale, des politiques de santé et de la recherche, de l'évaluation et/ou de l'amélioration de la qualité des services de santé (que nous désignons collectivement comme « recherche » dans ce document).

Ces intervenants et dépositaires de données comprenaient des représentants de la formation des médecins, du permis d'exercice des médecins, de la certification des médecins, de la réglementation des médecins, de l'accréditation des programmes de formation et des organisations de défense des médecins et des médecins en formation impliquées dans la trajectoire de développement professionnel médical. En tenant compte des avantages et des risques potentiels associés à la recherche inter-institutionnelle fondée sur les données, la série a abouti à un accord sur les principes fondamentaux et un ensemble de 3 recommandations pour l'utilisation éthique des données sur l'éducation à des fins d'enseignement médical, de politique de santé et recherche sur les services de santé au Canada.

Principes Fondamentaux

Confidentialité et anonymat	Consentement éclairé	Autonomie
Non-discrimination	Collecte de données appropriée	Objectifs de recherche appropriés

Recommandations

La recherche inter-institutionnelle en éducation fondée sur les données doit inclure:

1. La mise en place d'un organe de gouvernance qui comprend des représentants bien informés de toutes les parties prenantes concernées, garantissant les objectifs de recherche appropriés, la transparence, le consentement éclairé, le respect de la souveraineté des données, l'examen éthique et la surveillance proactive.
2. L'implication d'un service de gestion des données fiable qui permet les aspects techniques et procéduraux du partage des données.
3. Un engagement à gérer les données selon les meilleures pratiques qui incluent l'élaboration de stratégies institutionnelles, le développement de plans de gestion des données et l'observation des dépôts de données.

UN APPEL À L'ACTION

Présentement, les études qui utilisent des données longitudinales sur la formation médicale ne sont pas la norme. Cependant, le besoin de ces études continuera de croître à mesure que les intervenants demandent de plus en plus de preuves de la responsabilité et de l'impact de l'éducation médicale. Le statu quo de la gouvernance et de la surveillance de ce type de bourses d'études médicales comporte de nombreux risques. Les capacités et les modèles nécessaires à une gouvernance sensible et réceptive doivent être développés dans le contexte de la formation médicale et de ses parties prenantes. De plus, sans une réflexion et une action réfléchies sur les risques et les avantages du partage de données, notre domaine prendra du retard et créera peut-être les conditions d'un travail irréfléchi et potentiellement dangereux. Par conséquent, le but de cette déclaration de consensus est de servir de cadre pour relever les défis difficiles de la gouvernance, de l'éthique, de la responsabilité et des problèmes techniques alors que nous entrons dans un état où la recherche sur les « grandes données » devient la norme, et non l'exception.

Ces valeurs, ces principes et ces recommandations ne sont pas destinés à servir de politique de gouvernance officielle pour la recherche inter-institutionnelle en éducation fondée sur les données au Canada. Nous reconnaissons que chaque gestionnaire de données et chaque établissement d'enseignement médical au Canada opère dans le contexte d'une réalité organisationnelle et juridictionnelle unique qui dicte les processus de gestion des données qu'ils doivent prendre en compte dans la recherche de preuves de recherche qui soutiennent leurs missions. Nous espérons, cependant, que les établissements et les chercheurs reconnaissent la pertinence de ces valeurs, de ces principes et de ces recommandations en les adoptant alors qu'ils collaborent à des recherches axées sur les données dans le secteur de la formation médicale. Nous reconnaissons que la mise en application sera un défi. Nous pensons que des principes fermes peuvent guider une mise en œuvre flexible dans chaque contexte unique.

Nous demandons aux dirigeants de la communauté canadienne de la formation médicale de i) considérer et d'adopter ces principes et ii) d'étudier la mise en œuvre de ces principes et les recommandations qui le permettent, le cas échéant.

L'INITIATIVE DE RENFORCEMENT DES CAPACITÉS DE GESTION DES DONNÉES DE RECHERCHE

La formation médicale est un domaine interdisciplinaire qui fait progresser la théorie et la pratique de la formation des futurs médecins. Au cours de la formation médicale, les programmes d'enseignement génèrent de grandes quantités de données sur les étudiants. Cela comprend les informations sur les candidatures et les admissions, les performances d'évaluation (par exemple, les tests), les registres de présence, les rapports de professionnalisme, la participation à des activités parascolaires, les résultats des examens de licence, les notes de certification, les enregistrements des activités cliniques, les lieux de pratique après l'obtention du diplôme et diverses autres mesures provenant de nombreuses sources institutionnelles telles que les écoles médicales et professionnelles de la santé, les autorités de délivrance des permis et de certification, les organismes de réglementation et les agences provinciales de ressources humaines en santé. En règle générale, ces données sont collectées mais ne sont pas utilisées au-delà de leurs objectifs prédéfinis (par exemple, pour déterminer la compétence). Il est encore plus rare que ces données soient volontairement liées entre les périodes de formation et de pratique.

Dans le cadre d'un projet de Connexion parrainé par le CRSH, intitulé : « Initiative de renforcement des capacités de gestion des données de recherche : Éducation des professions de la santé », les intervenants et les responsables des données de l'ensemble de la communauté canadienne de l'éducation médicale ont participé à une série de 5 ateliers successifs de synthèse des connaissances sur invitation seulement qui visaient à faire avancer la réflexion collective et à établir un consensus en ce qui concerne la collecte et l'utilisation des données sur l'éducation des professions de la santé. En particulier, la série a réuni des personnes des domaines de l'enseignement supérieur, de la bibliothéconomie et des sciences des données, et des services d'épidémiologie/santé, qui ont fait des progrès considérables en matière de qualité des données et de normes de données pour partager leurs points de vue et leurs expériences à l'attention de ceux qui œuvrent dans le domaine de l'éducation médicale au Canada. Le but de cette série était de déterminer les principes fondamentaux pour l'utilisation appropriée des données sur l'éducation aux fins de la formation médicale, des politiques de santé et de la recherche sur les services de santé. Cela impliquait de prendre en considération le large éventail de tâches et de conséquences potentielles pouvant accompagner le partage de données inter-institutionnel pour la bourse.

Les objectifs du projet de connexion étaient de :

1. Faciliter la circulation des connaissances relatives à l'éthique des données et aux politiques entre les chercheurs, les administrateurs et les parties prenantes au sein de la communauté de l'enseignement médical.
2. Soutenir la mobilisation des connaissances sur la gouvernance, la qualité et la gestion des données pour la recherche entre les personnes des domaines de la bibliothéconomie, de l'épidémiologie/des services de santé et de l'enseignement supérieur qui ont développé des solutions à des défis similaires, et la communauté de l'enseignement médical.

Le résultat escompté du projet Connexion était une déclaration de consensus sur l'utilisation éthique des données administratives sur la formation pour la recherche en éducation médicale au Canada qui était conforme à la politique proposée des trois organismes sur la gestion des données de recherche.

NOTRE APPROCHE

Les objectifs, les perspectives, les avantages, les risques, les valeurs, les recommandations et les principes de gouvernance décrits dans ce document ont été dérivés d'une revue de la littérature; informés grâce à des consultations préalables à l'atelier avec des établissements d'enseignement médical canadiens qui recueillent, génèrent et conservent des données sur l'éducation, des intervenants en éducation médicale et des chefs de file de la recherche en science des données sur l'éducation (voir l'annexe A pour un aperçu des résultats des consultations préalables à l'atelier); reflétés par notre propre expérience en tant que scientifiques de l'éducation travaillant au sein de la communauté de recherche en éducation médicale au Canada; et avancés grâce au projet de connexion parrainé par le CRSH avec un groupe national d'intervenants pertinents. Cet événement était organisé sous la forme d'une série de 5 ateliers successifs de synthèse des connaissances qui visaient à faire avancer la réflexion collective et à établir un consensus. Les ateliers ont eu lieu les 25 mars, 1er avril, 7 avril, 14 avril et 21 avril 2021. Les ateliers duraient chacun 3 heures. À chaque séance, les organisateurs de la réunion ont attiré l'attention des participants sur le projet de politique des trois organismes sur la gestion des données de recherche, l'état de la synthèse émergente et les activités de l'atelier.

Chaque atelier a été complété par une présentation sur invitation d'un expert pertinent dans les domaines de l'éthique, de l'enseignement supérieur, de la bibliothéconomie et des services d'épidémiologie/santé; chacun d'entre eux avait une perspective unique sur les défis éthiques du partage de données sur l'éducation inter-institutionnelle ou avait fait des progrès considérables dans le partage de données dans leur propre domaine. Les conférenciers invités du projet Connection étaient la Dre Victoria McKinnon, représentant les perspectives des participants (Université McMaster, OMSA, CFMS); Dre Lisa Schwartz, une bioéthicienne de premier plan (Université McMaster); Dr Daniel Corral, un chercheur expérimenté des données massives dans l'enseignement supérieur (Université de Toronto); Dre Elizabeth Wenghofer qui a mené plusieurs projets de données longitudinales reliant l'éducation aux résultats des systèmes cliniques et de santé (Université Laurentienne); et Jeff Gardner et Nick Rochlin (Portage Network; Université de la Colombie-Britannique) qui ont fourni des points de vue d'experts sur les problèmes techniques et logistiques liés au partage de données. Ces points de vue d'experts ont servi à stimuler la réflexion des participants sur les caractéristiques fondamentales et les fondements d'un accord concernant l'utilisation inter-institutionnelle des données sur l'éducation aux fins de la recherche en éducation médicale; y compris le large éventail de tâches et de conséquences potentielles qui peuvent accompagner le partage de données inter-institutionnel pour la recherche. Dans chaque atelier, les participants se sont engagés dans des activités de groupe animées spécifiques à la session qui étaient orientées vers l'élaboration et le raffinement de l'accord de consensus. Dans chaque cas, ces activités comprenaient des exercices de réflexion en petits groupes qui ont été suivis d'un dialogue animé en grand groupe. Collectivement, ces activités visaient à promouvoir l'empathie pour toutes les parties prenantes, à favoriser une génération d'idées et à offrir des opportunités de productivité orientées vers l'élaboration d'accords. Notamment, chaque session a été suivie par un groupe d'observateurs, composé de représentants des organisations de défense des droits des médecins stagiaires concernés. Ce groupe a réfléchi sur les progrès de l'atelier en partageant leurs points de vue pendant une activité de table ronde dédiée à la fin de chaque session. Cette réflexion a permis d'intégrer les orientations éthiques relatives à la perspective du stagiaire dans le document final. Les ordres du jour des ateliers, les détails sur les conférenciers et leurs présentations, le cahier d'exercices du projet de connexion et d'autres ressources clés qui ont soutenu les sessions peuvent être consultés et examinés sur www.dataconnection.ca.

Au total, les ateliers ont été suivis par environ 50 personnes, représentant une grande variété d'institutions et d'intérêts, avec des représentants de l'Association des facultés de médecine du Canada, du Conseil médical du Canada, du Collège des médecins de famille du Canada, du Collège royal des médecins et chirurgiens du Canada, de la Fédération des ordres des médecins du Canada, du Registre canadien de l'éducation post-MD, de la Fédération canadienne des étudiants en médecine, de l'Association des étudiants noirs en médecine du Canada et des communautés de recherche en éducation médicale et en éducation médicale associées aux 17 établissements de formation médicale du Canada (annexe B). Tout au long de chaque session, la discussion a été menée de telle sorte que le groupe a examiné les perspectives, les contraintes et les opportunités uniques inhérentes au partage de données sur l'enseignement inter-institutionnel et ce que cela signifie pour la recherche en enseignement médical au Canada.

Ce document reflète notre synthèse critique de ces discussions. Au cours des ateliers, nous avons sollicité des idées et des commentaires sur des propositions pertinentes, et compilé des notes prises sur le vif concernant les réponses et les discussions des participants. À la fin de chaque session, l'équipe d'animateurs a examiné et classé ces notes en fonction des thèmes, des questions et des problèmes pertinents; un processus descriptif visant à saisir ce qui a été dit par les participants à l'atelier. Au fur et à mesure que les notes étaient regroupées, des idées consensuelles ont été identifiées, tout comme les idées pour lesquelles le consensus restait controversé, ambigu ou peu clair. Ces idées ont ensuite été utilisées par les animateurs de la série pour orienter les discussions au cours des sessions suivantes; de telle sorte qu'au cours de la série Connexion, les objectifs, les avantages, les risques, les valeurs et les recommandations suivants ont été définis.

Avant sa présentation finale, les versions préliminaires de ce document ont été révisées en réponse aux commentaires gracieux fournis par les représentants de l'Association des facultés de médecine du Canada, du Collège des médecins de famille, du Collège royal des médecins et chirurgiens du Canada, de l'Association canadienne pour l'éducation médicale, de l'Association des étudiants noirs en médecine du Canada, de la Société de la médecine rurale du Canada, des médecins résidents du Canada, de la Fédération des organismes de réglementation médicale du Canada et de la Fédération canadienne des étudiants en médecine ainsi que par le Dr Jean-Michel Leduc (Université de Montréal) et la Dre Christina St-Onge (Université de Sherbrooke) (voir l'annexe C pour un aperçu schématique de la série d'ateliers).

OBJECTIFS ET PERSPECTIVES PARTAGÉS

Les divers intervenants et dépositaires de données dans la trajectoire de perfectionnement professionnel médical comprennent la formation des médecins, le permis d'exercice des médecins, la certification des médecins, la réglementation des médecins, l'accréditation des programmes de formation et la défense des intérêts des médecins et des médecins en formation. Bien que chacun représente sa propre fonction distincte, à travers la série d'ateliers, nous avons reconnu un objectif commun :

Des ressources humaines en santé auto-réglementées et un système de santé canadien efficace pour répondre constamment aux besoins de soins de santé en constante évolution de toutes les personnes qui y ont accès, ainsi qu'aux aspirations professionnelles des médecins qui le composent.

Les conversations du projet Connexion ont également souligné que les intervenants médicaux canadiens et les dépositaires de données partagent 3 perspectives principales :

- La formation médicale est un facteur important pour atteindre l'objectif susmentionné.
- Les données éducatives sous toutes leurs formes ont le potentiel de soutenir la génération de connaissances susceptibles d'améliorer les pratiques et les politiques éducatives. Le manque de partage de données peut empêcher une intervention rapide et significative à l'avancement de la pratique de l'éducation médicale.
- La recherche fondée sur les données a déjà contribué à l'adoption de politiques et de pratiques qui ont favorisé les préjugés et accentué les inégalités sociales; et ce potentiel existe dans le contexte actuel. Des perspectives sur l'éducation et la recherche en éducation dans ce domaine qui ne sont pas informées des inégalités historiques, qui ne tiennent pas compte des droits et des priorités des communautés marginalisées et qui ne sont pas conscientes de la valeur de l'autonomie peuvent causer des dommages.

Inhérente à ces perspectives est une tension entre l'équilibre entre le bien potentiel à tirer de la recherche inter-institutionnelle fondée sur les données et le préjudice que ce type de travail peut causer aux médecins en formation et aux médecins, leurs objectifs professionnels et la prestation des soins de santé dans de nombreuses collectivités au Canada.

AVANTAGES DE LA RECHERCHE INTER-INSTITUTIONNELLE SUR L'ÉDUCATION BASÉE SUR LES DONNÉES

La discussion sur le partage de données inter-institutionnelles pour la recherche en éducation est pertinente car de nombreuses avancées dans la science des données en éducation servent d'animateurs pour ce type de recherche : les établissements d'enseignement médical collectent et archivent désormais de grandes quantités de données numérisées sur l'enseignement médical; de plus en plus d'outils sont développés pour soutenir la collecte de données standardisées entre les institutions; et l'évolution de l'infrastructure des données et des méthodes de recherche permet désormais des analyses rapides et à grande échelle. Par conséquent, même s'il est toujours nécessaire d'établir une technologie de l'information, une infrastructure de données et une préparation organisationnelle appropriées, la communauté de recherche en éducation médicale est de plus en plus motivée et dynamique pour tirer parti de ce nouveau potentiel et de ses nombreux avantages.

À cet égard, les avantages identifiés de faire ce type de travail comprennent :

- Rehausser les efforts d'amélioration de la qualité de l'éducation et d'évaluation. Grâce à la recherche sur le partage des données, les institutions peuvent parvenir à une meilleure compréhension de leur propre contexte unique.
- Faciliter la réussite des missions de responsabilité sociale. Grâce à la recherche sur le partage des données, nous pouvons arriver à une meilleure compréhension de la relation entre les politiques et pratiques d'éducation et l'efficacité du système de santé au niveau communautaire et les résultats pour la santé.
- Améliorer la qualité de la recherche en éducation. Les permis de partage de données créent plus d'opportunités pour la recherche sous forme de bourses qui mènent à une théorie de l'éducation qui est largement généralisable. Ce travail peut concerner de nombreux domaines de la recherche en éducation : évaluation, admissions, professionnalisme, compétence, développement des compétences, etc.

RISQUES DE LA RECHERCHE INTER-INSTITUTIONNELLE SUR L'ÉDUCATION BASÉE SUR LES DONNÉES

Malgré les moyens techniques et une description des avantages potentiels, le partage de données inter-institutionnelles pour la recherche en éducation n'est pas encore devenu une pratique courante. Il y a plusieurs raisons pour cela et les obstacles au progrès incluent les problèmes techniques relatifs à l'équivalence des données ; les contraintes temporelles et les implications en termes de ressources de l'examen éthique et de la documentation juridique ; un manque de structures de gouvernance adéquatement représentatives ; et un manque de normes de gestion des données. À l'origine de ces obstacles se trouve également une préoccupation répandue selon laquelle la poursuite du bénéfice présumé verra les données utilisées d'une manière qui expose les individus et les institutions à des risques, qui concernent à la fois les aspects procéduraux du partage des données ainsi que la manière dont les résultats potentiels de la recherche sur le partage de données peuvent avoir un impact sur les individus et les communautés.

Facteurs de risque

- La recherche axée sur les données peut être utilisée pour discriminer les médecins en formation et les médecins (dans plusieurs contextes) sur la base de caractéristiques d'identité personnelle ou sociale.
- La recherche axée sur les données favorise une culture de surveillance qui peut avoir une incidence sur les médecins en formation et sur le bien-être et le sentiment de sécurité des médecins. C'est particulièrement le cas pour les individus issus de groupes historiquement défavorisés par les méthodes de surveillance.
- Les résultats de la recherche basée sur les données peuvent être erronés et/ou mal interprétés de manière potentiellement préjudiciable. Cela peut se produire à cause d'une inexactitude ou d'une insuffisance des données; d'un manque de considération pour le contexte approprié; et/ou d'un constat inapproprié des limites de la recherche.
- Le partage des données fait place à de nouvelles formes de comparaison inter-institutionnelle dont certaines auront un impact négatif sur certaines institutions en ce qui concerne la perception du public. La recherche sur le partage des données peut également créer des opportunités pour un examen minutieux résultant d'une trop grande insistance sur des mesures ou des formes de données particulières.

Risques liés à la procédure

Le développement des ensembles de données inter-institutionnels peut contrevenir aux engagements institutionnels visant à :

- Confidentialité et anonymat. Les menaces à la sécurité des données (c'est-à-dire les fuites, les piratages, le vol) peuvent comprendre la sécurité et la confidentialité individuelles et institutionnelles, et peuvent entraîner une perte de données ou une désanonymisation. La confidentialité et l'anonymat peuvent aussi être compromises par des processus d'identification résiduelle (c'est-à-dire l'identification d'individus par l'incorporation d'une plus grande multitude de données). Les engagements institutionnels en matière de protection de la confidentialité et de l'anonymat tiennent souvent compte des principes d'équité en matière d'information présentés dans le Code modèle pour la protection des renseignements personnels de l'Association canadienne de normalisation (<https://laws-lois.justice.gc.ca/fra/lois/p-8.6/page-7.html>).

- Consentement éclairé. Les individus et les institutions ne sont pas toujours conscients de la manière dont leurs données peuvent être utilisées. Les processus d'obtention du consentement ou de détermination de l'autorité sur l'utilisation des données peuvent ne pas couvrir toutes les applications futures possibles.
- Promouvoir la souveraineté des données ou l'idée que les données doivent être soumises aux structures de gouvernance au sein du contexte ou de la communauté dans laquelle elles sont collectées. Cela peut faire référence à des conceptualisations de flux de données transnationales; cependant, dans le contexte actuel, c'est plus pertinent pour les notions d'autonomie autochtone vis-à-vis des États post-coloniaux et les désirs d'autres communautés pour une autonomie similaire en matière de données. La collecte, l'analyse et l'interprétation des données ne sont jamais vraiment « neutres » ou objectives et une dissociation des utilisations des données par rapport aux communautés qui ont fourni les données peut amplifier l'oppression et l'exclusion.

VALEURS PERTINENTES

L'événement Connexion a mis en évidence qu'une préoccupation clé pour tous les participants est l'établissement et le maintien de la confiance entre les parties prenantes et les chercheurs. La confiance est considérée comme un important facteur pour assurer la sécurité des individus, des communautés et des institutions. Au cours de nos conversations, nous avons défini ces valeurs qui, selon les intervenants de la formation médicale et les dépositaires de données, sont fondamentales pour établir cette confiance :

- *Consentement éclairé* : La collecte et le partage de données sur un individu doivent avoir lieu en sa connaissance. Cette valeur signifie de fournir aux personnes concernées un avis approprié sur l'utilisation des données, une vue transparente des pratiques de recherche et la faculté de choisir d'accepter.
- *Collecte de données appropriée* : Le suivi est un élément principal de la responsabilité institutionnelle et fondamental pour les pratiques d'évaluation de l'éducation. Par conséquent, les médecins en formation et les médecins comprennent que leurs progrès seront assujettis à la collecte de données. Cependant, la collecte de données sur les médecins en formation et les médecins devrait se limiter aux paramètres qui soutiennent les missions principales de l'institution concernée.
- *Objectifs de recherche appropriés*. La poursuite de la recherche en éducation fondée sur les données de manière opportuniste et/ou simplement dans le but d'exploiter l'infrastructure disponible et/ou d'améliorer la productivité de la recherche augmente les possibilités que les données puissent être mal utilisées, mal interprétées ou mal gérées (c'est-à-dire en fonction des risques liés à la procédure et aux résultats). À cet égard, la poursuite d'innovations opportunistes en matière de données peut nuire à la résolution de problèmes courants et solidement établis. Cela peut également contribuer à une culture de productivité de la recherche qui n'entraîne pas de changements significatifs dans les pratiques ou les politiques éducatives.
- *Confidentialité et anonymat* : Les individus doivent rester anonymes ou inconnus dans les ensembles de données de recherche. À cet égard, la quantité d'informations collectées et partagées doit être réduite au minimum à celle qui est requise pour l'objectif particulier. Presque toutes les institutions et juridictions ont mis en place des politiques pour protéger la vie privée des individus. La recherche axée sur les données doit respecter et incorporer les politiques et les pratiques pertinentes pour les institutions et les juridictions impliquées dans le projet.
- *Autonomie* : La recherche basée sur les données ne doit pas être utilisée pour limiter les options ou les possibilités de développement de ceux qui apprennent. L'analyse prédictive peut être utilisée pour restreindre la capacité d'un individu à gérer sa propre éducation, en guidant ou en poussant les gens dans certaines directions. La recherche axée sur les données ne doit pas influencer les participants ou les communautés d'une manière qui limite leur capacité à poursuivre des choix réfléchis et délibérés.
- *Non-discrimination* : Les données doivent être utilisées de manière à ne pas discriminer ou faire durer les inégalités. Les analyses qui font des prédictions sur les individus sur la base de constellations de leurs informations peuvent faire durer les préjugés, le profilage et la discrimination tout en accentuant la stratification sociale.

RECOMMANDATIONS

Avec ce qui précède, nous offrons les recommandations suivantes. Notamment, notre objectif ici n'est pas d'établir une politique, une procédure ou une approche spécifique pour le partage de données. L'intention est plutôt de fournir des éclaircissements à toutes les parties sur la façon dont le partage inter-institutionnel des données dans l'enseignement médical à diverses fins devrait être conceptualisé. Ceci est conçu comme un cadre d'un ensemble fondamental de recommandations, de principes et de perspectives spécifiques qui devraient être envisagés et appliqués à toute activité de partage de données en éducation médicale. À cet égard, nous reconnaissons que le contexte de la question peut avoir une influence considérable sur la façon dont les éléments suivants sont adoptés. Ces idées peuvent sous-tendre les relations formelles et servir de base à des accords de partage de données qui servent de médiateur à la recherche inter-institutionnelle en éducation; ou peuvent guider le développement d'énoncés de réflexivité qui servent à rendre transparentes les façons dont les données ont été gérées et la recherche a été menée. On espère, quel que soit le contexte, que les chercheurs et ceux qui gèrent et supervisent les données sur l'éducation soient attentifs à ces recommandations et à ces principes lorsqu'ils se lancent dans leur enquête.

Notamment, les recommandations portent sur le partage des données existantes, mais visent également à être complètes en abordant les nouvelles données recueillies au cours de la mise en œuvre de la formation médicale (voir l'annexe D pour les notes concernant la collecte et le partage des données d'identité sociale).

Recommandations

Nous recommandons ce qui suit :

- Un ou plusieurs organismes de gouvernance doivent être établis conformément aux principes de gouvernance qui suivent. Le rôle de cet organisme serait de superviser les accords de partage de données, d'éduquer la communauté sur le partage de données sûr et approprié, et d'examiner et de vérifier les activités universitaires de partage de données pertinentes. Cet organisme de gouvernance doit inclure une représentation des points de vue des apprenants, des médecins praticiens, des gestionnaires de données, des régulateurs, des établissements d'enseignement, des chercheurs et des membres de la communauté, y compris des représentants bien informés des communautés marginalisées et sous-représentées, en fonction des données partagées et des questions de recherche poursuivies. Lorsque les ressources ou d'autres contraintes empêchent une représentation cohérente, des mécanismes doivent être mis en place pour consulter les parties prenantes concernées selon les besoins. De plus, cet organisme devrait assurer la fonction d'examen éthique supplémentaire des projets de partage de données au-delà de l'examen local et, par conséquent, peut aussi inclure des experts en la matière concernant la législation sur la gestion des données et la confidentialité. L'organisme de gouvernance doit être légitimement habilité à adopter et à superviser le partage des données qui s'aligne sur les règles, les politiques et les lois des juridictions souveraines dans lesquelles les données à partager ont été collectées.

- Un ou des services de gestion de données fiables doivent être établis pour permettre les aspects techniques et procéduraux du partage de données dans la recherche en éducation médicale ainsi que pour appliquer les procédures et politiques approuvées par l'organisme de gouvernance. Ce service peut administrer des accords de partage de données de novo, assumer les accords existants lorsque la gouvernance et la législation l'approuvent, et travailler avec les gestionnaires de données existants le cas échéant. Le service serait également responsable de l'assurance de la qualité des données, de la communication à l'organisme de gouvernance de la sécurité des données et du respect des protocoles. Ce service permettrait aussi d'assurer et de superviser, si nécessaire, la désagrégation et le retour des données, ainsi que l'oubli des données (c'est-à-dire l'expiration des données). Un tel service peut ou non avoir un rôle dans l'archivage des données.
- Les données doivent être gérées conformément aux normes alignées sur les meilleures pratiques, y compris celles décrites dans la politique proposée des trois conseils sur la gestion des données de recherche (c.-à-d. considérations relatives aux stratégies institutionnelles articulées, aux plans de gestion des données et aux dépôts de données). En plus de ces pratiques, nous recommandons que les données partagées à des fins d'analyse minimisent l'utilisation d'identifiants individuels. Des identifiants spéciaux doivent être utilisés pour l'accouplement et le nombre maximal d'identifiants doit être supprimé avant les analyses. Il peut y avoir des situations dans lesquelles les marqueurs institutionnels peuvent aussi être supprimés des ensembles de données avant l'analyse. Il peut y avoir des situations dans lesquelles la gouvernance permet l'inclusion de ces informations; selon la nature de la question de recherche, l'assentiment des parties prenantes participantes et les risques que cela représente. Lorsque c'est possible, les techniques de gestion, de traitement et de transformation des données devraient se concentrer sur et/ou accentuer la variance expliquée par les politiques, les procédures et les systèmes d'éducation plutôt que sur celle qui se rapporte aux individus dans les données; dans la mesure où la question de recherche l'exige.

Principes de gouvernance

Nous proposons que le partage de données sur la formation médicale entre les établissements à des fins de bourses d'études et d'activités universitaires soit régi par les principes suivants :

1. La mise sur pied *a priori* d'objectifs, de justification et de méthodologie clairs pour le partage des données. Bien que ceux-ci puissent être vastes et couvrir diverses activités savantes, ils doivent être clairement définis et prendre en considération les avantages et les inconvénients. Il se peut qu'il ne soit pas possible de définir les questions de recherche spécifiques pour chaque programme de travail soutenu par le partage de données. Cependant, la gouvernance doit garantir que les types d'associations/liens et les méthodologies utilisés - y compris les dispositions pour la protection de la confidentialité, de la vie privée et de la minimisation des dommages - sont clairement définis.

2. La transparence dans la communication de ces avantages et de ces inconvénients, ces risques et ces protections constitue le fondement du consentement éclairé. Le consentement éclairé de ceux qui génèrent des données dans les systèmes éducatifs, y compris les candidats, les étudiants, ainsi que les enseignants, le personnel et les organisations, est un aspect non négociable de la recherche. Ce consentement éclairé doit être négocié par des processus de mét-consentement (c'est-à-dire concernant différents types potentiels d'utilisation secondaire tels que déterminés par un accord de partage de données pertinent) qui doivent être « limités dans le temps » (c'est-à-dire donnés pour une certaine période pour les utilisations consenties), puis résilié, renouvelé ou renégocié. Ce consentement éclairé est particulièrement important pour le partage de données impliquant des caractéristiques d'identité sociale. Le consentement limité dans le temps offre la possibilité de respecter l'autonomie des générateurs de données/participants avec de nouvelles informations sur les risques/avantages résultant des nouvelles données créées par le couplage de données ainsi que de confirmer la validité continue des données dans le système. Ce processus devrait être une responsabilité partagée entre les chercheurs, les services de partage de données et les organes directeurs responsables de la responsabilité du partage des données dans le but de minimiser le fardeau du consentement pour les participants.
3. Le partage des données crée des risques de processus et de résultats uniques qui nécessitent une gouvernance et une responsabilité spéciales qui doivent être sensibles aux risques et aux avantages ainsi qu'une connaissance des processus de recherche. Ces risques ne sont pas répartis uniformément entre toutes les populations et toutes les communautés. Certaines communautés, en particulier celles qui ont des antécédents d'exclusion, devraient faire l'objet d'une attention particulière en ce qui a trait aux risques.
4. Conformément au principe 3, les utilisations des données devraient être co-gouvernées par toutes les parties prenantes concernées. Cela signifie que les structures de gouvernance doivent inclure une représentation experte ou bien informée des institutions, des communautés et d'autres groupes qui détiennent une souveraineté, une gérance ou d'autres relations significatives avec les données. C'est particulièrement important pour les communautés qui ont été historiquement marginalisées ou opprimées par rapport aux pratiques sociales et scientifiques. Les membres avertis de ces communautés devraient être particulièrement invités à faciliter la création de la gouvernance.
5. Les structures de gouvernance pour le partage des données doivent être guidées par le respect de la « souveraineté des données », c'est-à-dire les règles/politiques/lois au sein de la communauté, de la juridiction ou de l'institution dans laquelle elles sont collectées. Les représentants concernés doivent être identifiés a priori pour tout partage de données et impliqués dans la création et/ou la révision et l'amélioration de l'accord de partage de données. À cet égard, les chercheurs doivent s'engager en permanence avec l'organe de gouvernance tout au long du processus de recherche, y compris pendant l'interprétation des résultats et avant la diffusion des résultats pour s'assurer que la recherche présente des avantages significatifs et/ou ne perpétue pas les dommages pour les communautés, les juridictions et les institutions affectées.
6. Les structures de gouvernance devraient examiner les utilisations proposées des données par rapport aux critères de respect des valeurs fondamentales (c.-à-d. autonomie, consentement éclairé, confidentialité et anonymat, collecte de données appropriée, objectif approprié, non-discrimination), du mérite scientifique et du potentiel d'impact bénéfique. Quelle que soit l'activité prévue du partage de données (par exemple, l'amélioration de la qualité, l'évaluation de programme, la recherche, etc.), les risques sont les mêmes et toujours pertinents. Ainsi, il doit incomber aux chercheurs de s'engager dans les niveaux appropriés d'examen éthique, quelle que soit la catégorie d'activité universitaire.

7. Les structures de gouvernance doivent être pro-actives et vigilantes pour superviser et garantir l'utilisation des données conformément aux principes, aux processus et aux politiques spécifiques établis dans les accords de partage de données et les pratiques de recherche approuvées. La gouvernance doit promouvoir des plans de gestion des données alignés sur les meilleures pratiques, y compris les politiques actuelles et futures adoptées par les trois conseils sur la gestion des données de recherche.

ANNEXES

A : Consultations et thèmes de la série de pré-ateliers

B : Liste des participants

C : Aperçu de la série de Connexion de données éducatives

D : Concernant la collecte et le partage des données d'identité sociale

ANNEXE A : CONSULTATIONS ET THÈMES DE LA SÉRIE DE PRÉ-ATELIERS

En préparation pour la série Connexion CRSH, nous avons mené une série de consultations avec des dépositaires de données, des chercheurs scientifiques et des groupes d'adhérents étudiants de partout au Canada.

- Dépositaires : Association des facultés de médecine du Canada, Conseil médical du Canada, Fédération des ordres des médecins du Canada, Collège des médecins de famille du Canada, Collège royal des médecins et chirurgiens du Canada, Association canadienne de protection médicale
- Chercheurs scientifiques : Christina St-Onge (Université de Sherbrooke), Jean-Michel Leduc (Université de Montréal), Suzanne Archie (Université McMaster), Onye Nnorom (Université de Toronto), LLana James (Université Queen's), Saad Chahine (Université Queen's), Douglas Archibald (Université d'Ottawa)
- Groupes d'adhérents étudiants : Association des étudiants noirs en médecine du Canada, Association des médecins autochtones du Canada, Fédération canadienne des étudiants en médecine

Notre objectif lors de ces consultations était d'identifier les principales préoccupations relatives à la recherche inter-institutionnelle en éducation médicale fondée sur les données. Ici, nous mettons en évidence notre synthèse de ces conversations; condensant une série de thèmes qui ont informé nos conversations.

Thème 1 : Équilibrer les mandats de responsabilité sociale

Nos consultants ont régulièrement invoqué leurs mandats de responsabilité sociale institutionnelle/organisationnelle comme raison de participer ou de ne pas participer au partage de données inter-institutionnel à des fins de recherche. Ce faisant, le groupe met collectivement en évidence deux malaises qui doivent être équilibrés lors de la prise de décisions pour participer à ces initiatives.

- a. Le premier malaise concerne le conflit qui peut exister au sein des missions sociales propres d'une organisation. Par exemple, plusieurs groupes reconnaissent le pouvoir de la recherche fondée sur les données comme un outil efficace d'amélioration de la qualité pour faire avancer les efforts vers une meilleure planification des effectifs de la santé; l'amélioration de l'éducation et des résultats de la formation; l'amélioration de l'équité, de la diversité et de l'inclusion dans la profession médicale; et/ou des résultats sociaux et sanitaires améliorés. À cet égard, il y a un sentiment qu'il est impératif d'utiliser les données à ces fins. Cependant, plusieurs des mêmes institutions soulignent leur responsabilité sociale de protéger leurs membres/constituants contre des préjugés injustifiés et de respecter leur confidentialité et leur anonymat. À cet égard, il y a un sentiment parallèle qu'il y a une obligation de ne pas participer au partage de données à des fins de recherche.
- b. Le deuxième malaise concerne les conflits qui peuvent exister entre les divers mandats des différentes institutions et ce que cela peut signifier sur la façon dont la recherche est conceptualisée, les ensembles de données sont utilisés et les résultats de la recherche sont exploités. Par exemple, les facultés de médecine s'intéressent à la recherche en éducation fondée sur les données pour mettre en évidence les domaines d'amélioration des programmes ou des politiques. Les régulateurs, quant à eux, ont intérêt à identifier les risques potentiels pour les communautés de patients. En faisant cette comparaison, nous pouvons voir que les deux groupes peuvent adopter des positions différentes sur les questions posées, les données recherchées et les innovations recherchées.

Thème 2 : Les enjeux logistiques ou techniques du partage de données pour la recherche

Nos consultants ont régulièrement souligné les défis techniques du partage de données comme une raison de participer ou de ne pas participer à la recherche inter-institutionnelle axée sur les données. Plusieurs défis ont été mis en évidence sous ce thème, notamment ceux relatifs aux processus d'harmonisation et de normalisation des champs de données entre les organisations, à la création d'accords efficaces de partage de données entre les institutions et à la conduite d'activités de recherche dans des délais raisonnables.

Thème transversal : L'ajout de données basées sur l'identité

Au cours de nos conversations, un sujet singulier a émergé comme étant pertinent à la fois pour le Thème 1 et le Thème 2, qui délimitait deux types de données : les données d'admission et d'évaluation actuellement détenues et utilisées par les différents gestionnaires de données; et les données d'identité sociale (par exemple, concernant la race, le sexe, l'origine ethnique) qui ne sont pas actuellement collectées mais qui, selon la plupart des coordonnateurs et des organisations, devraient être collectées dans le but d'effectuer des recherches qui produisent des preuves qui informent les progrès dans les missions de responsabilité sociale

En ce qui concerne le Thème 1, il était clair que nos consultants ont reconnu que les implications de l'échange de données sont touchées lorsque l'on change d'orientation pour inclure les données d'identité sociale. Ces données étaient clairement considérées comme essentielles pour l'avancement de mandats efficaces en matière d'équité, de diversité et d'inclusion. Cependant, beaucoup ont noté que la simple collecte et l'analyse de ces données sont insuffisantes. À cet égard, la tension entre l'usage et l'abus était très répandue. Il existe des inquiétudes (présentes et historiques) selon lesquelles les données d'identité sociale peuvent être utilisées pour nuire, stigmatiser et ne pas améliorer la vie des personnes représentées. À cet égard, nos consultations ont souligné l'importance de comprendre le contexte dans lequel ces données sont situées et la nécessité de reconnaître que le potentiel des avantages de la recherche prosociale à être compensé par des résultats oppressifs existe à toutes les étapes du continuum de la recherche axée sur les données - à partir de la question de recherche jusqu'à l'interprétation des résultats. Quant au Thème 2, nos discussions ont également souligné que l'inclusion de données d'identité sociale a des implications pour les aspects techniques du partage de données. En particulier, l'inclusion de données d'identité sociale met davantage l'accent sur la composition des structures de gouvernance des données qui assurent la médiation des approbations de recherche et des accords de partage de données. Ces données augmentent également la difficulté de maintenir l'anonymat des étudiants au sein d'un ensemble de données.

ANNEXE B : LISTE DES PARTICIPANTS

- Doug Archibald (Université d'Ottawa)
- Cassandra Barber (Université de Maastricht)
- Geoffrey Barnum (CAPER)
- Ariel Barrientos (BMSAC, Université de Toronto)
- Ilona Bartman (CMC)
- Brenton Button (EMNO)
- Alice Cavanagh (Université McMaster)
- Erin Cameron (EMNO)
- Saad Chahine (Université Queen's)
- Shen Chu (Université McMaster)
- Daniel Corral (Université de Toronto)
- Annabelle Cumyn (Université de Sherbrooke)
- Gurm Dhugga (UBC)
- Holly Fleming (EMNO)
- Jeffrey Gardner (Réseau Portage, UBC)
- Qi Guo (CMC)
- Mark Hanson (Université de Toronto)
- Deena Hamza (Université de l'Alberta)
- Tanya Horsley (CRMCC)
- Rebekah Jacques (Université Western)
- Christopher Jones (Université de Toronto)
- Eleni Katsoulas (Université Queen's)
- Jonathan Kimball (AFMC)
- Dragan Klujic (CMFC)
- Jean-Michel Leduc (Université de Montréal)
- Hana Lee (Université de Toronto)
- Rachelle Lee-Krueger (Université d'Ottawa)
- Fleur-Ange Lefebvre (FOMC)
- Henry Li (FEMC, Université du Manitoba)
- Cathy MacLean (Université de la Saskatchewan)
- Laura McEwen (Université Queen's)
- Victoria McKinnon (FEMC, AÉMO, Université McMaster)
- Kestrel McNeill (Université McMaster)
- Jeffrey Moat (Pallium Canada)
- Susan Moffatt-Bruce (CRMCC)
- Sandra Monteiro (Université McMaster)
- Ivy Oandasan (CMFC, Université de Toronto)
- Gbolahan Olarewaju (BMSAC, UBC)
- Ghislaine Pilot-Attema (EMNO)
- Nick Rochlin (Réseau Portage, UBC)
- Shelley Ross (ACÉM; Université de l'Alberta)
- Lisa Schwartz (Université McMaster)
- Bolade Shipeolu (BMSAC, Université de Toronto)
- Steve Slade (CMFC)
- Christina St-Onge (Université de Sherbrooke)
- Brent Thoma (Université de la Saskatchewan)
- Kimberley Thomas (BMSAC, UBC)
- Elizabeth Wenghofer (Université Laurentienne)
- Brianne Wood (EMNO)
- Meredith Young (Université McGill)
- Alaa Youssef (Université de Toronto)

Remarque : Tous les participants n'ont pas assisté à toutes les sessions.

ANNEXE C : APERÇU DE LA SÉRIE DE CONNEXION DE DONNÉES ÉDUCATIVES

Consultations de pré-série

En préparation pour la série Connexion CRSH, nous avons mené une série de consultations avec des dépositaires de données, des chercheurs scientifiques et des groupes d'étudiants de partout au Canada. De ces conversations, nous avons défini une série de thèmes qui ont éclairé nos conversations.



Série d'ateliers sur la Connexion de données

Cette série de 5 ateliers en ligne a permis de partager les connaissances relatives à l'éthique et à la politique des données parmi les chercheurs et les parties prenantes et a contribué à la création de la déclaration de consensus.

Atelier 1 :
Considérations éthiques et perspectives des étudiants

Atelier 2 :
Expériences de l'enseignement supérieur

Atelier 3 :
Expériences de la recherche sur les services de santé

Atelier 4 :
Expériences en bibliothéconomie

Atelier 5 : Examen de la nouvelle déclaration de consensus



Consultations post-séries

ANNEXE D : CONSIDÉRANT LA COLLECTE ET LE PARTAGE DES DONNÉES D'IDENTITÉ SOCIALE

Les conversations sur la série Connexion ont mis en évidence une suggestion particulièrement convaincante concernant la recherche en éducation qui nécessite la collecte et le partage de données sur l'identité sociale; en particulier : *que les risques de partage de données inter-institutionnels sont trop élevés et que nous ne devrions pas nous engager dans ce type de recherche en éducation.* C'est une idée qui a été classée comme devant maintenir le statu quo.

À cet égard, nous reconnaissons que les récents appels à la collecte et à l'analyse de données sur l'identité sociale encadrent ce type d'activité à l'appui de l'enquête qui peut fournir des preuves qui informent les avancées des théories de la responsabilité sociale et les résultats des missions d'équité, de diversité et d'inclusion. Bien que plusieurs institutions et groupes de défense s'engagent dans une certaine forme de collecte de données démographiques sociales, généralement sous la forme d'enquêtes anonymes qui cèdent la place à des ensembles de données agrégées anonymes, ces données ne sont actuellement pas recueillies d'une manière normalisée qui lie directement ces caractéristiques aux médecins en formation et aux médecins.

L'événement Connexion a mis en évidence que l'incertitude entre l'utilisation potentielle et la mauvaise utilisation de ces données est très répandue. Il y a des inquiétudes actuelles et anciennes selon lesquelles les données d'identité sociale peuvent être utilisées pour nuire, stigmatiser et ne pas améliorer la vie des personnes représentées par les données. En conséquence, il a été convenu que nous devons reconnaître le potentiel que les avantages de la recherche prosociale peuvent être compensés par des résultats oppressifs à toutes les étapes du continuum de la recherche axée sur les données - de la question de recherche à l'interprétation des résultats, et que l'inclusion de données d'identité sociale augmente la probabilité qu'un individu puisse être identifié dans un ensemble de données.

À cet égard, nous reconnaissons que les valeurs mentionnées ci-dessus doivent être particulièrement prises en compte dans la conduite de toute recherche en éducation qui implique la collecte et/ou le partage de données d'identité sociale. Il est impératif que ce type d'utilisation des données appuie la recherche qui aborde et ne perpétue pas l'inégalité sociale. Il doit reconnaître que ces données augmentent la probabilité que les individus puissent être identifiés au sein des ensembles de données et que les interprétations des analyses prédictives peuvent servir à diminuer l'autonomie de l'étudiant. Dans le cas où les objectifs, les justifications et les méthodes de recherche exprimés a priori ne sont pas en mesure de rendre compte de ces risques potentiels, nous recommandons alors que les données d'identité sociale relatives aux médecins en formation et aux médecins ne soient pas collectées, partagées ou analysées.