

# Connaissances acquises sur l'utilisation des données de mobilité en santé publique

## PETIT DÉPART POUR LES MÉGADONNÉES

**P**eu après l'écllosion de la pandémie de COVID-19, l'Agence de la santé publique du Canada (ASPC/l'Agence) a fait appel à d'innovantes sources de données et à l'expertise du secteur privé pour guider la riposte à une crise de santé publique qui prenait rapidement de l'ampleur. Elle a engagé plusieurs fournisseurs de mégadonnées et, de concert avec d'autres services gouvernementaux, a commencé à tenter de comprendre les données dont elle disposait.

Au départ, elle ne possédait pas d'équipe spécialisée dans l'étude de ce type de données, destinées à être utilisées dans le domaine de la santé publique. Les premiers travaux ont plutôt été effectués par une petite équipe, soutenue par sa direction, très intéressée à engager le dialogue avec les épidémiologistes de l'ASPC. L'expertise diversifiée, acquise au fil d'un travail de collaboration réalisé avec les responsables de la surveillance de COVID-19 et leurs partenaires, a largement participé à l'orientation initiale du dossier.

L'équipe Discovery a été créée sur la base de ce travail, pour développer une capacité d'utilisation de données non traditionnelles au sein de l'Agence. Elle a obtenu un accès sous licence aux données agrégées et dépersonnalisées d'appareils mobiles. Cela a marqué le début de travaux approfondis sur la valeur et les utilisations possibles des mégadonnées, en appui à la riposte à la pandémie.

## Contenu

Cette étude de cas montre comment, dans une situation donnée, l'équipe a utilisé des données d'appareils mobiles pour contribuer à une intervention de santé publique. Ses observations et ses expériences peuvent aider les décideurs, les gestionnaires, les épidémiologistes, les analystes de politiques et d'autres acteurs de la santé publique à se préparer à travailler avec des mégadonnées.



Centre de collaboration nationale  
des maladies infectieuses

National Collaborating Centre  
for Infectious Diseases

Pour l'équipe, travailler avec des nouvelles données dans un contexte de pandémie très tendu a représenté un véritable défi. Elle a dû bien sûr en apprendre beaucoup sur les données et sur la manière dont celles-ci pouvaient être utilisées dans le domaine de la santé publique, et ce, dans le contexte de la pandémie de COVID-19. Mais les travaux entrepris ont également permis de stimuler l'innovation, la croissance et le développement de nouvelles compétences au sein de l'équipe.

Cette étude de cas relate l'expérience d'une équipe et ne doit pas être lue comme un parti-pris en faveur des mégadonnées ou comme un guide sur l'exploitation des données d'appareils mobiles. Elle peut toutefois aider à comprendre les options, les possibilités et certaines voies prometteuses pour la santé publique dans ce domaine en progression rapide.

## À PROPOS DES DONNÉES DE MOBILITÉ

L'utilisation généralisée des téléphones cellulaires a créé de grandes quantités de données sur le comportement humain, reconnues comme offrant une riche source d'informations sur les mouvements de population, avec une variété d'applications dans différents domaines, dont celui de la santé publique (1). Pour saisir le potentiel de ces données pour la santé publique, il faut d'abord comprendre certains des types de données et leurs caractéristiques.

Il existe deux types courants de données de mobilité qui peuvent être tirées des appareils mobiles : les données des opérateurs et celles obtenues par approches participatives. *Les données des opérateurs* sont collectées par les fournisseurs de services cellulaires à partir des connexions régulières des appareils aux tours cellulaires avoisinantes ou lors de la transmission d'appels, de messages ou de courriels. *Les données obtenues par approches participatives* fournissent des renseignements géographiques obtenus grâce à la localisation d'un appareil mobile recevant des signaux du système mondial de géolocalisation (GPS) et collectés à partir d'applications participantes figurant sur les appareils des utilisateurs, lorsque ces derniers ont activé le service de localisation. Certaines caractéristiques distinguent ces deux types de données l'un de l'autre (voir le tableau à la page 3) et tous deux présentent des avantages et des inconvénients, en fonction de l'usage qui en est fait.

## L'équipe Discovery

L'équipe Discovery est une unité de l'Agence de la santé publique du Canada qui soutient l'analyse des données et renforce les capacités d'exploitation de données nouvelles.

Elle apporte une perspective axée sur la santé publique et une expertise politique et technique à la prestation de plusieurs services spécialisés, notamment le soutien à l'acquisition de données, l'évaluation de nouvelles sources de données et la conception et le développement de systèmes qui collectent, stockent et intègrent les données dans des outils de visualisation et d'analyse.

Les sources de données non traditionnelles que l'équipe exploite pour soutenir les priorités en matière de santé publique s'appuient sur la recherche, la consultation et les exigences en matière de protection de la vie privée. Elles orientent l'évaluation des besoins, les analyses de rentabilité et la participation des parties prenantes et des utilisateurs.

L'équipe Discovery promeut l'utilisation de données et méthodologies nouvelles au sein de l'Agence, en formulant des recommandations à l'équipe de direction.

L'équipe Discovery a examiné ces deux types de données en fonction des types d'analyses auxquelles elles peuvent servir. Au niveau de la population, les deux types de données permettent de réaliser des analyses similaires, portant notamment sur les tendances générales en matière de déplacement, sur la connectivité, en indiquant le nombre de déplacements entre différentes localités, et sur les points d'intérêt, en révélant les habitudes des gens liées à certains lieux, tels que les visites à l'épicerie ou à l'hôpital.

Bien que le potentiel analytique des données des opérateurs et de celles obtenues par approches participatives soit similaire, leur spécificité géospatiale diffère, ce qui en limite l'analyse. Dans un échantillon, les données collectées avec une approche participative reposent sur des coordonnées géographiques pour déterminer l'emplacement des appareils mobiles. Grâce à des méthodes d'agrégation complexes, les points de données de localisation dépersonnalisés, mais précis des données

recueillies par approches participatives peuvent signaler des changements dans les tendances et les schémas de mouvement régionaux. Les données des opérateurs reposent sur la densité des tours de téléphonie cellulaire, laquelle diffère d'une région à l'autre (les centres urbains ont plus de tours que les zones rurales et isolées). Dans les zones peu peuplées où les tours de téléphonie mobile sont peu nombreuses, les données des opérateurs, dans un échantillon de données, n'indiquent qu'approximativement la position des appareils. Cela peut donc limiter l'analyse des mouvements de population relatifs aux points d'intérêt de ces zones.

L'équipe de Discovery a également examiné ses données du point de vue de l'éthique en matière de santé publique.

Les valeurs et les principes d'éthique en matière de santé publique régissent la manière dont l'Agence s'acquitte de son mandat de promouvoir et de protéger la santé des Canadiens, y compris la façon dont elle tient compte de la confidentialité et de la sécurité des données. Dès le départ, l'équipe a procédé à une évaluation critique des données en fonction de considérations d'ordre éthique et a reçu les conseils d'experts internes en matière de droit et de protection de la vie privée. Par ailleurs, les fournisseurs de données sous contrat avec l'ASPC ont leurs propres pratiques et adhèrent à des politiques qui protègent les renseignements personnels des Canadiens. L'article de BlueDot Inc. (voir page 4) présente le point de vue de l'un de ces fournisseurs sur les considérations relatives à la confidentialité et à la sécurité des données.<sup>1</sup>

## Types de données de la mobilité et leurs caractéristiques

TYPE DE DONNÉES	Précision géospatiale	Représentation de la population	Collecte des données	Temps d'accès aux données	Types d'analyse
<b>Données par approches participatives (à partir d'applications)</b>	Positionnement de la latitude et de la longitude avec une précision quasi exacte dans toutes les provinces et tous les territoires.	La taille de l'échantillon est tributaire du nombre d'utilisateurs de l'application.  La taille de l'échantillon varie quotidiennement ou hebdomadairement.	Activée lorsque les utilisateurs ont choisi de partager leur position sur des applications (par exemple, restaurants, applications d'achat).  Position saisie lorsque l'utilisateur utilise et active l'application.	Données accessibles en temps quasi réel (délai d'environ une à deux semaines)	Tendances générales des déplacements  Connectivité entre l'origine et la destination  Schémas de mouvement par rapport à des points d'intérêt définis  Rassemblements
<b>Données des opérateurs (fournisseurs de services cellulaires)</b>	Tributaire de la densité des tours de téléphonie cellulaire. Un plus grand nombre de tours permet une plus haute précision.  Varie entre les provinces et les territoires (par exemple, composition urbaine ou rurale différente).	La taille de l'échantillon est tributaire du nombre d'abonnés du fournisseur.  Fluctuations rares et faibles	Activée régulièrement par les tours de téléphonie cellulaire pour les appareils recevant un signal et par les événements liés aux téléphones cellulaires (c'est-à-dire les appels, les textes, les courriels)	Données accessibles en temps réel (délai de deux à sept jours environ)	Tendances générales des déplacements  Connectivité entre l'origine et la destination  Schémas de mouvement par rapport à des points d'intérêt définis

1. La mention de contenus émanant de BlueDot Inc. ne signifie pas que l'Agence de la santé publique du Canada les approuve ou les recommande.

# POINT DE MIRE

## CONFIDENTIALITÉ ET SÉCURITÉ DES DONNÉES DE MOBILITÉ

*Contribution de BlueDot Inc.*

BlueDot accorde une importance capitale à la protection de la vie privée et à la sécurité des données, ce qui exige une vigilance de tous les instants dans toutes les relations qu'elle entretient avec ses clients et ses fournisseurs de données. Les considérations relatives à la protection de la vie privée dont BlueDot tient compte lors de l'acquisition et de la préparation des données sont riches d'enseignements pour quiconque travaille avec des données similaires.

### Éléments à prendre en compte lors de la sélection d'un fournisseur

La protection de la vie privée commence par une sélection rigoureuse des fournisseurs de données. En ce qui concerne les mesures de protection de la vie privée, les fournisseurs de données ne sont pas tous égaux et certains sont basés dans des pays où les lois sur la protection de la vie privée sont moins strictes que celles du Canada ou que celles de leaders mondiaux en la matière, comme la Californie.<sup>2</sup> Bien que de nombreux facteurs influencent la sélection d'un fournisseur, les considérations relatives à la protection de la vie privée qui devraient influencer cette décision comprennent :

- le respect des lois sur la protection de la vie privée reconnues par l'industrie, s'appliquant également à toutes les données ;
- des audits, des certifications et des pratiques en matière de sécurité transparents ;
- des options de refus et de consentement claires pour les utilisateurs d'appareils ou d'applications individuels ;
- des restrictions et des réglementations claires sur l'utilisation des données fournies.

Certaines limitations importantes à l'utilisation des données comprennent l'obligation d'éliminer les données à la fin d'un contrat, les seuils minimaux pour l'agrégation de données, la gestion des processus de dépersonnalisation effectuée par le fournisseur et les restrictions interdisant toute tentative de réidentification des données.

### Mesures additionnelles pour la protection des renseignements personnels

À partir de ces considérations initiales, d'autres mesures devraient être prises pour adopter des pratiques exemplaires en matière de protection de la vie privée. Lorsque BlueDot prépare des données pour ses clients, elle vérifie que les données transmises par les fournisseurs ne renferment aucun renseignement sur les propriétaires des appareils et que ces fournisseurs utilisent plutôt un identifiant anonyme généré de manière aléatoire. Les données ne comportent aucun renseignement au sujet d'un appareil autre que l'heure et la géolocalisation associées à l'identifiant anonymisé. De même, les applications fournissant des données de localisation

2. BlueDot Inc. adhère à la *Loi sur la protection des renseignements personnels et les documents électroniques* (LPRPDE) du Canada ainsi qu'aux politiques de pointe de l'industrie, notamment la *California Consumer Privacy Act* et le *Règlement général sur la protection des données* de l'Union européenne.

sont limitées à celles qui ont un objectif raisonnable de collecte d'informations de localisation, qui ont l'autorisation de recevoir ces données (par exemple, les applications pour les restaurants, les achats, etc.) et qui exigent que les utilisateurs donnent leur accord, par le biais des paramètres de localisation de leur appareil ou de l'application elle-même.

Dans le cadre du contrat passé par BlueDot avec les fournisseurs de données et afin de préserver la confidentialité et la représentativité statistique, les résultats créés par BlueDot sont agrégés par unités géographiques (par exemple, région sanitaire) ou unités de temps (par exemple, rapports quotidiens, hebdomadaires ou mensuels). En outre, les données sont supprimées lorsque le nombre d'appareils est trop faible pour garantir la validité et l'exactitude des statistiques. Cette mesure supplémentaire renforce la protection de la vie privée et doit être prise chaque fois que l'on travaille avec ces données.

### Le respect des normes en matière de protection de la vie privée profite à la santé publique

Le maintien de normes strictes en matière de protection de la vie privée ne relève pas seulement d'une question de principe ; il est essentiel à l'obtention de résultats valables susceptibles de faire avancer la santé publique. Les épidémies sont influencées par les comportements des populations dans le temps, selon qu'elles sont plus ou moins mobiles, selon les fluctuations de la connectivité entre les lieux et selon la mesure dans laquelle les décisions politiques, la communication en matière de santé publique et la progression des épidémies se reflètent dans les changements de comportement. C'est pourquoi l'application de normes strictes en matière d'agrégation des données ne protège pas seulement la vie privée des utilisateurs d'appareils, mais correspond au niveau de rigueur nécessaire pour effectuer des analyses de santé publique qui peuvent contribuer à la réflexion sur les politiques à mener.

## BlueDot Inc.

BlueDot est une société torontoise spécialisée dans l'application de l'intelligence artificielle (IA) à la surveillance des maladies infectieuses et à l'évaluation des risques. Fondée en 2013, BlueDot a développé une plateforme d'information sur les épidémies en appui aux interventions rapides en maladies infectieuses.

BlueDot se procure divers ensembles de données provenant de sources publiques et commerciales et les analyse dans le but de détecter, dans le monde entier, les signaux d'épidémies dès leurs premiers stades, de prévoir les tendances de propagation de ces épidémies par le biais d'un réseau mondial de vols et d'appuyer les interventions locales qui en atténuent les conséquences sanitaires, économiques et sociales.

La pandémie de COVID-19 a mené à la création d'applications remarquables de la technologie de BlueDot en matière de santé publique, notamment la détection précoce de l'épidémie de Wuhan en décembre 2019 et la prévision précise de la dispersion mondiale de la COVID-19, comme le relève une étude évaluée par des pairs (2).

## EXPLORATION DES DONNÉES, DÉFINITION DES OBJECTIFS

Pour commencer à travailler avec une source de données nouvelles, l'équipe de l'ASPC a dû se familiariser avec les données et acquérir l'expertise nécessaire pour effectuer les analyses appropriées. Elle a commencé par se demander : «*Quelles sont les questions de santé publique auxquelles ces données pourraient répondre?*». Ses objectifs d'utilisation des données ont peu à peu pris forme en étudiant des cas d'utilisation possibles et en apprenant à mieux connaître les limites des données.

Dans un contexte de pandémie mondiale, l'ASPC a commencé par fonder des espoirs sur l'information qu'il serait possible d'obtenir à partir des données de mobilité, tout en se montrant prudente quant aux considérations éthiques que cela pourrait poser. Même si les données ne permettaient pas de savoir si les individus respectaient les consignes de santé publique (distanciation physique de deux mètres ou isolement), il était possible d'évaluer l'adhésion générale de la population aux mesures de confinement. Les données de mobilité dépersonnalisées et agrégées acquises par l'ASPC ne refléteraient que les mouvements de masse des populations sans fournir de renseignements sur les individus.

Les analystes de l'équipe ont pris le temps d'apprendre à travailler avec des ensembles de données volumineux et de plus en plus nombreux, et à interpréter leur représentation des mouvements de population. Au début du projet, l'équipe a examiné les types de données de mobilité provenant des opérateurs et par approches participatives et s'est demandé : «*En ce qui concerne les déplacements des Canadiens, les deux ensembles de données racontent-ils une histoire similaire?*» Leur analyse a permis de comparer le même indicateur (mouvement général) dans le temps, aux niveaux national, provincial ou territorial, et au niveau des régions sanitaires. Les analystes ont ensuite posé la question suivante : «*Les deux ensembles de données montrent-ils des mouvements de population qui augmentent et diminuent en même temps?*» Ces comparaisons ont révélé des tendances similaires et ont permis de valider les données, confortant l'équipe dans ce qu'elle pouvait dire sur les mouvements de population.

Dès le début de la pandémie, l'équipe de l'ASPC avait espéré que les données de mobilité aideraient à prédire où les épidémies étaient susceptibles de se produire et où les ressources devaient être déployées pour en

prévenir la propagation. Cependant, à mesure que le taux et la complexité de la transmission communautaire augmentaient rapidement, les données aux fins de prévision et de prévention ont perdu de leur utilité.

Les besoins et les priorités en matière d'analyse évoluaient au fur et à mesure que l'épidémiologie de COVID-19 se modifiait. En communiquant régulièrement avec les responsables de la surveillance de la COVID-19, l'équipe a pu recueillir des avis qui l'ont aidée à orienter ses analyses. En s'appuyant sur les informations recueillies auprès des acteurs de la lutte contre la pandémie, l'équipe a déterminé qu'il serait plus judicieux d'utiliser les données de mobilité pour comprendre les modèles et les tendances historiques des mouvements de population au cours de la pandémie en relation avec les mesures de santé publique adoptées dans les provinces ou au niveau fédéral.

## L'ÉVALUATION DES DONNÉES DE MOBILITÉ

Les données de mobilité sont très différentes des sources de données de santé publique traditionnelles et offrent de nouvelles perspectives, tout en posant de nouveaux défis. Après avoir travaillé sur les données d'appareils mobiles lors de la riposte à la pandémie, l'équipe de l'ASPC s'est penchée sur les forces et les limites de ces données. Les points saillants de leur évaluation, laquelle peut aider à comprendre les points forts et les limites de données similaires du point de vue d'un utilisateur en santé publique, sont présentés ci-dessous.

### Les forces et le potentiel des données

#### Une mesure valide

Afin de mieux comprendre la validité de ses données, l'équipe de l'ASPC a consulté à la fois des études et de la littérature grise, selon lesquelles les données d'appareils mobiles représentent une mesure raisonnablement précise des déplacements de population (3). Les conclusions de divers spécialistes de l'Agence, d'autres services gouvernementaux et d'organisations non gouvernementales ont également contribué à des évaluations positives de la validité des données. On estime que l'immense volume de données de mobilité améliore la validité parce que l'effet du biais d'échantillonnage est réduit. Les données sont également moins affectées par la désirabilité sociale et les erreurs de mémoire que d'autres sources, telles que les données d'enquête.

## Informations en temps quasi réel

Les données d'appareils mobiles fournissent des informations en temps quasi réel, ce que l'équipe de l'ASPC juge utile pour la prise de décision. Elles sont généralement traitées avec un décalage d'une semaine, alors que les données des opérateurs sont disponibles dès le lendemain. Un aperçu en temps quasi réel de la situation actuelle peut aider la santé publique à préparer une réponse rapide en cas d'épidémie de maladie infectieuse.

## Utilisation dans les travaux de modélisation

Les données de mobilité sont utiles pour certains types de modélisation de maladies que l'équipe de l'ASPC a étudiées. Les modèles qui combinent des informations sur la façon dont la population se déplace avec d'autres ensembles de données, tels que les admissions à l'hôpital, les vaccinations et les taux de maladie, peuvent aider la santé publique à valider des hypothèses ou à prédire où les maladies infectieuses se propageront.

## Coût-efficience

En raison des volumes importants de données obtenues, les données d'appareils mobiles peuvent s'avérer plus efficaces que les données d'enquête, dont les échantillons sont relativement petits, en particulier pour les analyses de données à l'échelle nationale que l'ASPC a été en mesure de réaliser en interne. De plus, le nombre d'heures-personnes nécessaires à l'élaboration, à la mise en œuvre et à l'analyse des données de mobilité est peu élevé.

## Limites et défis des données

### Risque perçu

La perception du public selon laquelle le gouvernement pourrait se servir des données des téléphones portables pour suivre les déplacements des gens fait craindre à la santé publique une attention médiatique négative et pourrait décourager l'utilisation des données de mobilité. L'équipe de l'ASPC explique aux parties prenantes que la valeur pour la santé publique de ces données très volumineuses ne réside pas dans les données individuelles, mais bien dans les mouvements de population, et plus particulièrement dans l'évolution de ces déplacements dans le temps. Malgré cela, l'utilisation secondaire des mégadonnées n'est pas bien comprise et la lutte contre la désinformation relève du défi. En outre, il est difficile, en raison des méthodes complexes appliquées pour

dépersonnaliser et agréger les données en vue de leur utilisation dans le domaine de la santé publique, d'expliquer aux personnes ayant des connaissances techniques limitées les mesures de protection des renseignements personnels adoptées à l'égard de leurs données.

## Dotation de ressources

Bien que plus économiques que d'autres méthodes d'enquête, l'achat et l'exploitation de mégadonnées peuvent s'avérer coûteux, même s'il est possible d'accéder à certaines sources gratuites. Les coûts sont susceptibles d'inclure l'ajout de personnel possédant des compétences spécialisées pour travailler avec les données. Par ailleurs, les volumes importants et toujours croissants de mégadonnées exigent une infrastructure informatique solide, comprenant des serveurs supplémentaires, des processeurs rapides, une plateforme en nuage et une grande capacité de stockage numérique. L'équipe de l'ASPC a insisté sur le fait que sans ces ressources, son analyse des données d'appareils mobiles aurait été inefficace, voire impossible à effectuer.

## Nécessité d'une expertise interne

Les organismes de santé publique peuvent manquer de personnel formé et aguerri dans le calcul et l'analyse de mégadonnées. Bien que des consultants du secteur privé puissent combler un manque et que la collaboration avec d'autres organisations soit essentielle, l'intégration d'un savoir-faire technique au sein de l'équipe comporte des avantages. L'ASPC a donc choisi d'embaucher du personnel possédant la formation et l'expérience nécessaires pour rejoindre une équipe chargée de ce travail. Cette expertise interne a permis de guider judicieusement l'analyse en tenant compte des perspectives de la santé publique, de veiller à ce que les règles d'éthique de la santé publique soient respectées et qu'une évaluation des risques par rapport aux avantages soit faite.

## Manque d'informations sociodémographiques

Les données de mobilité étant agrégées et dépersonnalisées, il n'est guère possible de les associer à des attributs sociodémographiques tels que l'âge, le sexe ou l'appartenance ethnique, ce qui ne permet pas à la santé publique de bien saisir les différences entre les schémas de mobilité au sein de la population. L'équipe a reconnu que cela limitait la disponibilité des données pouvant servir à élaborer des ripostes adaptées.

## Représentativité des sous-populations

Les données d'appareils mobiles sont affectées par des biais démographiques liés à leur utilisation. Les gens qui ne possèdent pas d'appareils ou qui s'en servent moins fréquemment sont sous-représentés dans ces données, alors que ceux qui ont plus d'un appareil ou qui utilisent plus fréquemment leur(s) téléphone(s) sont surreprésentés. Les données peuvent être moins utiles pour estimer la mobilité dans des régions possédant certaines caractéristiques sociodémographiques, telle une moyenne d'âge élevée. Les données de mobilité tendent également à être moins représentatives des populations éloignées et rurales, car pour la plupart de ces zones géographiques (par exemple, les Territoires du Nord-Ouest, le Yukon et le Nunavut), la taille des échantillons est plus petite. L'ASPC a également noté que les fournisseurs de données provenant d'approches participatives ne donnent aucune information sur les différentes applications à partir desquelles leurs données de localisation sont collectées. Lorsque les acheteurs de données ne peuvent pas savoir dans quelle mesure leur échantillon est représentatif de l'ensemble de la population ou de groupes ayant des préférences particulières en matière d'applications, cela peut constituer une autre source de biais.

## Difficultés de validation

En l'absence de norme de référence permettant de comparer les données, la capacité de validation des données de mobilité demeure limitée. L'équipe de l'ASPC a signalé qu'il n'est pas possible de vérifier que les personnes se sont déplacées; les données indiquent plutôt les mouvements humains. Ces données se distinguent des autres sources secondaires que la santé publique exploite et qui autorisent certaines vérifications; les données administratives sur la santé, par exemple, permettent de vérifier les dossiers d'hospitalisation ou les traitements qui ont été administrés.

## Établir des inférences causales

Enfin, l'équipe de l'ASPC reconnaît que l'une des principales limites des données de mobilité tient au fait qu'il n'est pas possible d'en tirer des inférences causales sur les mouvements de population, même si certaines corrélations peuvent être établies.

## EXPLOITER LES DONNÉES DE MOBILITÉ

L'équipe de l'ASPC a exploré plusieurs cas d'utilisation de données de mobilité pour éclairer les politiques liées à la pandémie. Les exemples suivants présentent un cas d'utilisation probant qui s'est avéré utile dans leur travail, un cas d'utilisation prometteur qui a été testé avec un sous-ensemble de données, et un troisième cas d'utilisation possible, qui pourrait être appliqué à l'avenir.

### Cas d'utilisation probant

#### Mesures de santé publique et mouvements de population

Les données de mobilité ont été très utiles pour évaluer les mesures de santé publique en fonction de leur impact sur les mouvements de population. Pendant la pandémie, les provinces et les territoires ont instauré diverses mesures ou politiques de santé publique à différents moments, en fonction de l'épidémiologie locale. *L'indice de rigueur* (voir l'encadré à la page 8) représente le niveau de réponse employé et peut être rapproché des données sur la COVID-19 afin de signaler les interventions susceptibles d'avoir des effets.

L'ASPC a pu comparer les effets du resserrement ou du relâchement de la rigueur des mesures de santé publique sur l'augmentation ou la diminution des mouvements de population qui en découlent (voir le graphique ci-dessous). Cela a permis de comprendre comment la population réagissait aux mesures de santé publique. Il a également été intéressant de constater que les changements associés aux déplacements, très nets au début de la pandémie, sont devenus moins perceptibles par la suite. Cela peut témoigner d'une certaine lassitude à l'égard de la pandémie ou d'une confiance accrue dans l'immunité, liée à l'administration de vaccins ou à un récent rétablissement à la suite d'une maladie.

L'indice de rigueur comporte certaines limites : il peut notamment être affecté par le nombre de grandes villes dans un territoire donné, car celles-ci ont tendance à imposer des restrictions plus sévères que les petites villes et les régions (4). De futures études pourraient envisager d'appliquer les données à des zones géographiques plus ciblées afin de modérer le poids des grandes villes. Différents éléments de l'indice de rigueur pourraient également être comparés aux données de mobilité afin de déterminer lesquelles des politiques individuelles sont les plus efficaces pour réduire les mouvements de population.

## L'INDICE DE RIGUEUR

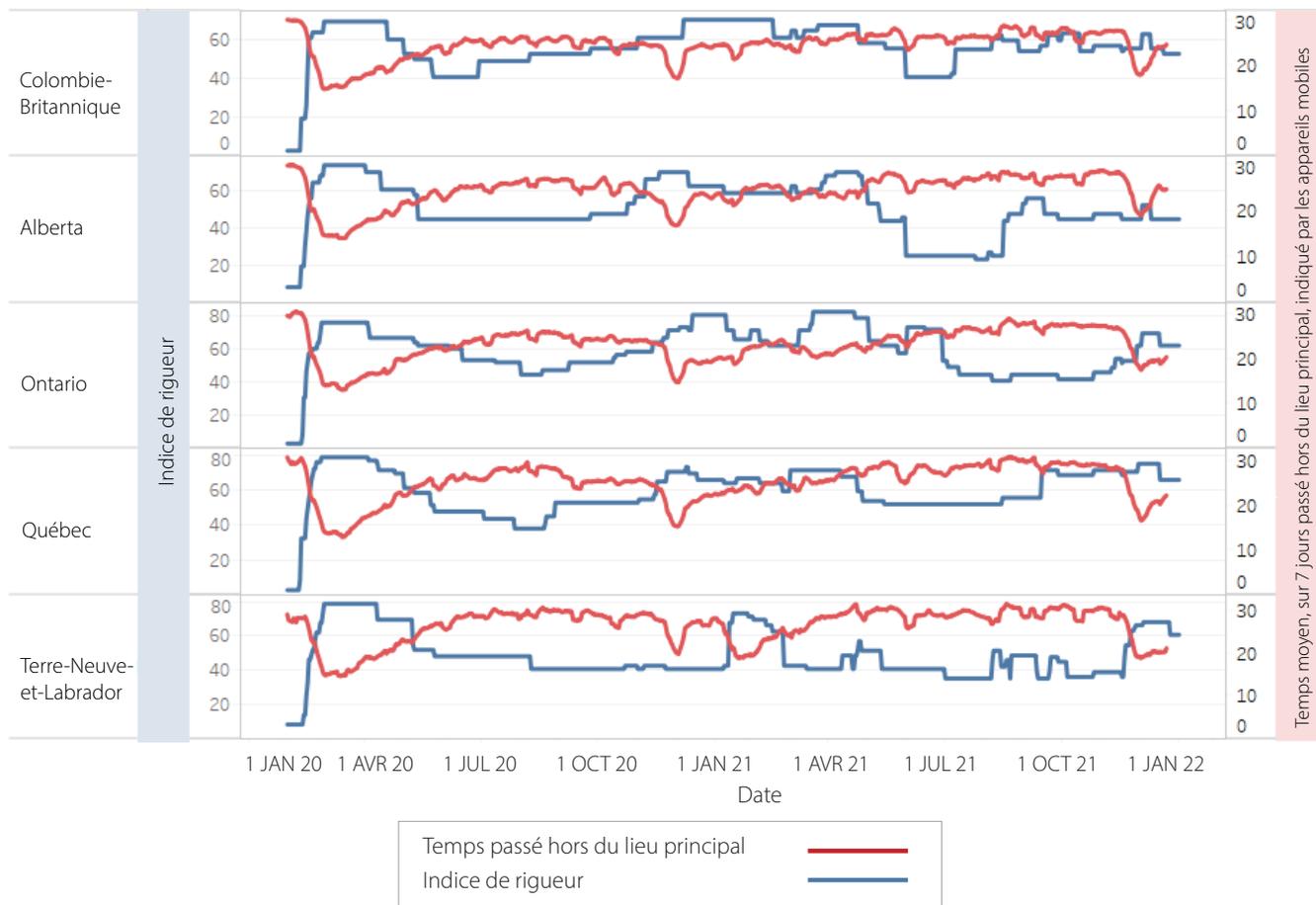
L'indice de rigueur des mesures sanitaires liées à la COVID-19, élaboré par l'Université d'Oxford, enregistre la rigueur des mesures et des politiques de santé publique qui limitent principalement le comportement des personnes (5).

L'indice est une mesure semi-quantitative qui associe des données provenant de neuf interventions de santé publique différentes : fermeture d'écoles, fermeture de lieux de travail, annulation d'événements publics, restrictions sur la taille des rassemblements, fermeture

des transports publics, obligation de rester chez soi, restrictions sur les mouvements intérieurs (c'est-à-dire à l'intérieur du pays ou d'une province ou d'un territoire), restrictions sur les voyages internationaux et campagnes d'information du public.

L'indice de rigueur correspond au niveau de réponse de la sous-région la plus stricte. Il ne mesure pas la pertinence ou l'efficacité d'une mesure et ne tient pas compte de la conformité ou de l'adhésion aux politiques et aux mesures.

### Temps passé hors du lieu principal comparé à l'indice de rigueur des mesures de santé publique dans certaines provinces canadiennes (janvier 2020 — décembre 2021)



**Source :** Indice de rigueur pour les provinces et territoires des mesures de santé publique de l'Agence de la santé publique du Canada (ASPC), élaboré par le Centre des maladies infectieuses d'origine alimentaire, environnementale et zoonotique (CMIOAEZ) de l'ASPC, et produit par le CMIOAEZ de septembre 2020 à février 2022, et par l'équipe de la Surveillance la COVID-19 du Centre de l'immunisation et des maladies respiratoires infectieuses (CIMRI) de l'ASPC de février 2022 à aujourd'hui.

## Un cas d'utilisation prometteur

### Déplacements dans les régions à haut risque

Les données de mobilité sont prometteuses pour détecter les régions sanitaires présentant un risque plus élevé de transmission de maladies dans les provinces et des territoires. Après avoir assoupli les restrictions en matière de santé publique, une province a connu une augmentation alarmante des cas de COVID-19. Cette augmentation rapide a entraîné la mise en place d'un état d'urgence dans la province et de « garde-fous » (c'est-à-dire l'interdiction de rassemblements ou des voyages dans d'autres régions) dans certaines régions afin de réduire la propagation de la COVID-19. Les mesures se sont traduites par une baisse de la mobilité dans la province et un examen plus approfondi de certaines zones de la province en question a révélé que celles dans lesquelles des « garde-fous » avaient été instaurés affichaient également une mobilité la plus faible. L'exemple montre que les données de mobilité peuvent être utiles pour suivre les résultats des mesures de santé publique sur les mouvements de population au niveau régional et pour comprendre les effets des mouvements au niveau local. Cette application pourrait également être étendue et développée pour aider à détecter les zones sanitaires à haut risque dans l'ensemble du pays.

### Cas d'utilisation potentiel

#### Points d'intérêt et évolution des comportements

Les travaux sur les données de mobilité de l'ASPC suggèrent que ces données pourraient donner des indications sur les tendances de comportement de la population, telles que la fréquentation de lieux particuliers (désignés par le terme « points d'intérêt »), comme les épiceries, les pharmacies, les magasins d'alcool, les hôpitaux, les établissements de soins de longue durée, ou les grands rassemblements organisés. Les tendances de la mobilité aux points d'intérêt peuvent être analysées sur des périodes données et comparées aux années précédentes. Ces informations peuvent servir à comprendre les tendances liées aux politiques de santé publique et à l'incidence ou à la prévalence des maladies ou des facteurs de risque. Par exemple, les visites dans les magasins d'alcool peuvent être un indicateur de l'évolution de la consommation d'alcool dans la population, avec des répercussions potentielles sur la santé psychologique. Associer les tendances en matière de mobilité pour tous

les points d'intérêt peut également servir d'indicateur de mobilité à l'échelle de la population.

### Réflexion sur la valeur

Ces cas d'utilisation et d'autres que l'ASPC a étudiés au cours de la pandémie de COVID-19 ont montré la valeur et le potentiel inexploité des mégadonnées. L'équipe a constaté que les informations tirées des données de mobilité étaient particulièrement utiles lorsqu'elle les combinait à des informations provenant d'autres sources, y compris des données de surveillance traditionnelles. Même si les données de mobilité n'ont pas directement servi à l'élaboration des politiques, elles se sont révélées un complément utile qui a permis de combler certains manques de connaissances. Si les travaux sur les données de mobilité que l'Agence a menés se sont concentrés sur la COVID-19, la littérature scientifique propose d'autres exemples de la façon dont la santé publique a fait usage de ce type de données (voir l'encadré à la page 10).

### CE QU'IL FAUT RETENIR

L'équipe Discovery a tiré des enseignements de ses travaux sur les données de mobilité — ce qui a bien fonctionné et ce qu'elle aurait pu faire différemment. Ces enseignements peuvent aider à comprendre à quoi s'attendre en travaillant sur des projets de mégadonnées et comment les planifier.

#### Obtenir des conseils d'experts, sur des questions de droit et d'éthique, par exemple

Les mégadonnées étant de nature complexe et sensible, la santé publique peut bénéficier de conseils professionnels pour procéder à un examen minutieux des implications de l'exploitation de ces données sur le plan de l'éthique, de la vie privée, du droit et de la sécurité. Par exemple, lors de l'achat de mégadonnées, il peut être utile de comprendre la différence entre les responsabilités légales et éthiques.

#### Informier le public sur les projets d'utilisation de mégadonnées

Il est important de faire preuve de transparence sur les projets d'utilisation de mégadonnées et d'expliquer précisément d'où celles-ci proviennent. Une annonce publique peut être faite dans un langage simple, décrire un cas d'utilisation prévu et en présenter les avantages espérés

## POINT DE MIRE

### APPLICATIONS DES DONNÉES DE MOBILITÉ

#### Dans la littérature de recherche sur la santé publique

Récemment, la santé publique a surtout exploité les données de mobilité dans le cadre de l'épidémie de COVID-19, mais la littérature scientifique distingue trois grands domaines d'utilisation :

#### Maladies infectieuses en santé publique

- Les données de mobilité sont utilisées pour modéliser les mouvements de population, la propagation des maladies, les points névralgiques et la connectivité entre les régions, illustrés par des cartes thermiques et des nœuds de transmission importants (6).
- Elles ont été exploitées avec d'autres données, notamment des données socio-économiques, pour comprendre par exemple la relation entre les niveaux de revenus régionaux et les mouvements de population (6,7).

#### Urbanisme

- Les données de mobilité ont le plus souvent été utilisées a posteriori, par exemple pour étudier les mouvements de population à long terme en fonction des systèmes de transport et des services sociaux présents en divers endroits (8,9).
- Le transport étant l'une des principales sources d'émissions de carbone, certaines recherches se sont penchées sur la mobilité liée aux modes de transport afin de déterminer comment les systèmes de transport pourraient être modifiés pour réduire leurs émissions de carbone et lutter contre le changement climatique (10).

#### Préparation aux catastrophes environnementales

- Les données de mobilité ont permis de déterminer où les ressources devraient être affectées en cas de situation d'urgence et comment améliorer l'efficacité de la répartition des ressources (11).

pour la santé publique. Il est également utile de préciser les options offertes au public, par exemple comment il est possible de refuser toute collecte de données.

### Garantir des ressources et des capacités suffisantes

Un plan de travail sur les mégadonnées exige une réflexion approfondie sur les ressources techniques et humaines à prévoir. Les travaux peuvent nécessiter un budget important pour permettre l'achat de nouveaux ordinateurs dotés d'une capacité de traitement rapide. Les projets de mégadonnées requièrent également du personnel disposant de certaines compétences techniques et d'une expérience de travail avec des mégadonnées. Idéalement, l'équipe sera pluridisciplinaire et comprendra des professionnels des programmes et des politiques de santé publique, des épidémiologistes, des spécialistes des données, des statisticiens et des modélisateurs.

### Travailler en réseau avec les premiers utilisateurs et apprendre d'eux

Travailler avec des données nouvelles et innovantes peut aboutir à des pistes de recherche improductives ou à des analyses sans intérêt. L'exploration de données non traditionnelles est imparfaite et réclame une ouverture à l'expérimentation qui peut ne pas répondre à vos besoins. Pour atténuer le risque de cette imprévisibilité, la santé publique peut investir du temps dans le travail collaboratif avec des mégadonnées. Discuter avec d'autres personnes ayant analysé des données similaires peut vous donner une bonne idée du potentiel de vos données — découvrez ce qui a bien fonctionné pour eux afin de vous économiser temps et efforts. Disposer d'un réseau pluridisciplinaire de collègues professionnels qui évaluent ouvertement et de manière critique le potentiel technique et la valeur de mégadonnées pour la santé publique peut s'avérer particulièrement utile dans ce secteur relativement nouveau que sont les applications de données pour la santé publique.

### Préparer un plan d'analyse

Bien que les analystes soient en mesure d'accomplir beaucoup de choses avec les mégadonnées, il est important qu'une analyse permette de répondre à des questions de santé publique importantes et qu'elle donne des informations que la santé publique peut exploiter. L'élaboration d'un plan d'analyse de données — une idée

de la manière dont celles-ci seront utilisées — peut s'avérer utile. Le plan peut s'appuyer sur ce que vous avez appris d'autres personnes sur ce qui fonctionne.

## **Soutenir les processus de gestion des données**

La taille volumineuse et l'expansion continue des ensembles de mégadonnées, de même que les conditions de licence susceptibles de leur être appliquées, rendent d'autant plus importante la mise en place d'un système de gouvernance des données. Il peut être utile de confier à un membre de l'équipe la responsabilité de gérer les données à chaque étape de leur cycle de vie, depuis l'acquisition et la mise en place des structures de stockage jusqu'à la mise en œuvre de protocoles d'accès aux données au sein de l'organisation, en passant par l'archivage et l'élimination appropriés des données. Les étapes préparatoires et la nécessité d'exercer toute la diligence requise dans la gestion des données sont souvent négligées, mais elles ont une valeur à long terme.

## **EN APPRENDRE PLUS**

Cette étude de cas évoque quelques-unes des expériences que l'ASPC a vécues, aux premiers jours de la pandémie de COVID-19, alors qu'elle découvrait et explorait les possibilités qu'offrent les données de mobilité. Bien que l'Agence ait depuis acquis de nouvelles connaissances, les premières leçons tirées de ces expériences peuvent aider d'autres organisations de santé publique à évaluer les possibilités et les difficultés qu'un travail similaire présente.

L'équipe de l'ASPC et ses collègues s'interrogent aujourd'hui sur ce qu'il est possible de faire de plus. Les données de mobilité pourraient-elles servir à décrire les niveaux d'activité prépandémiques et postpandémiques, les schémas de mouvement de la population et les modes de transport dans différentes régions, de même que leur lien avec les taux de maladies chroniques? Que pourraient nous apprendre les données de mobilité sur les schémas de mouvement humain par rapport à l'éventail des vecteurs de maladies infectieuses, pour que nous puissions être en mesure de prévoir la propagation des virus liée au changement climatique? Que pourrions-nous apprendre sur l'insécurité alimentaire en examinant la mobilité en fonction de l'emplacement des magasins d'alimentation au sein de diverses collectivités?

Au Canada, l'utilisation des données de mobilité et d'autres mégadonnées dans le domaine de la santé publique est

encore assez récente. Bien que la pandémie de COVID-19 ait stimulé l'innovation et la collaboration, dont une utilisation inattendue des mégadonnées, il reste encore beaucoup à faire.

## **REMERCIEMENTS**

Le CCNMI est hébergé par l'Université du Manitoba. Nous tenons à reconnaître que nous sommes rassemblés sur le territoire désigné par le Traité no 1, terre ancestrale des peuples anishinabé, cri, oji-cri, dakota et déné, et la terre natale de la nation métisse.

Le CCNMI remercie les membres de l'équipe Discovery d'avoir généreusement accepté de faire part de leurs observations et des leçons qu'ils ont tirées de l'utilisation de mégadonnées à des fins de santé publique, de leurs précieuses contributions et du temps considérable qu'ils ont consacré à l'élaboration de l'étude de cas.

Nous remercions également les équipes Epidemic Intelligence et Data Systems Development de BlueDot Inc. des informations détaillées sur la confidentialité et la sécurité des données qu'ils nous ont fournies, et de leurs conseils sur d'autres contenus techniques.

L'équipe de Discovery tient à remercier les nombreux épidémiologistes, chercheurs et experts techniques de l'Agence de la santé publique du Canada et du Centre de recherches sur les communications d'Innovation, Sciences et Développement économique Canada qui ont offert leur temps, leurs connaissances et leur énergie pour donner un sens à ces données massives, en pleine pandémie, alors que tout le monde devant composer avec de nombreuses priorités.

## RÉFÉRENCES

1. Oliver N, Matic A, Frias-Martinez E. Mobile network data for public health : opportunities and challenges. *Front. public health*. 7 août 2015 ; 3 : 189. <https://doi.org/10.3389/fpubh.2015.00189>
2. Bogoch II, Watts A, Thomas-Bachli A, Huber C, Kraemer MUG, Khan K. Potential for global spread of a novel coronavirus from China, *J. Travel Med.* Mars 2020 ; 27 (2). Consultable à l'adresse : taaa011, <https://doi.org/10.1093/jtm/taaa011>
3. Zhao C, Zeng A, et Yeung CH. Characteristics of human mobility patterns revealed by high-frequency cell-phone position data. *EPJ Data Science*. 2021 ; 10 (5):1-14.
4. Long JA, Malekzadeh M, Klar B, Martin G. Do regionally targeted lockdowns alter movement to non-lockdown regions? Evidence from Ontario, Canada. *Health and Place*. 2023; 79: 102668. <https://doi.org/10.1016/j.healthplace.2021.102668>
5. Mathieu E, Ritchie H, Rodés-Guirao L, Appel C, Giattino C, Hasell J, et coll. Coronavirus pandemic (COVID-19). Publié en ligne sur OurWorldInData.org. 2020. Consultable à l'adresse : <https://ourworldindata.org/coronavirus>
6. DIAL. Using mobile network operator data for COVID-19 response [Internet]. Digital Impact Alliance. 2020 Consultable à l'adresse : <https://digitalimpactalliance.org/using-mobile-network-operator-data-for-covid-19-response/>
7. Levin R, Chao DL, Wenger EA, Proctor JL. Insights into population behavior during the COVID-19 pandemic from cell phone mobility data and manifold learning [Internet]. *Nat Comput Sci*. 2021; 1:588–597. Consultable à l'adresse : <https://www.nature.com/articles/s43588-021-00125-9>
8. Cho SB, et Kim JY. Clustered embedding using deep learning to analyze urban mobility based on complex transportation data. *PloS one*. 2021; 16(4), e0249318. <https://doi.org/10.1371/journal.pone.0249318>
9. Nations Unies The state of mobile data for social good report [Internet]. United Nations Global Pulse; juin 2017. Consultable à l'adresse : [https://unglobalpulse.org/wp-content/uploads/2017/06/Mobile\\_Data\\_for\\_Social\\_Good\\_Report.pdf](https://unglobalpulse.org/wp-content/uploads/2017/06/Mobile_Data_for_Social_Good_Report.pdf)
10. Nieuwenhuijsen MJ. Urban and transport planning, environmental exposures and health-new concepts, methods and tools to improve health in cities. *Environmental health*. 2016; 15 (Suppl 1), S38. Consultable à l'adresse : <https://doi.org/10.1186/s12940-016-0108-1>
11. Wang Y, Li J, Zhao X, Feng G, et Luo XR. Using mobile phone data for emergency management: a systematic literature review. *Information systems frontiers*. 2020; 22(6): 1539–1559. Publication anticipée en ligne. <https://doi.org/10.1007/s10796-020-10057-w>



Centre de collaboration nationale  
des maladies infectieuses

National Collaborating Centre  
for Infectious Diseases

Salle L332A, Basic Medical Sciences Building  
745, avenue Bannatyne  
Faculté des sciences de la santé Rady  
Université du Manitoba, Winnipeg (Manitoba) R3E 0W2  
Tél. : (204) 318-2591  
Courriel : [ccnmi@umanitoba.ca](mailto:ccnmi@umanitoba.ca)  
Site Web : [www.ccnmi.ca](http://www.ccnmi.ca)



## À RETENIR

L'expérience de l'ASPC sur l'utilisation des mégadonnées a permis d'en tirer plusieurs enseignements :

- Obtenir des conseils d'experts, sur des questions de droit et d'éthique, par exemple
- Informer clairement le public des projets sur l'utilisation des mégadonnées
- Veiller à ce que votre équipe dispose de ressources et de capacités suffisantes
- Travailler en réseau avec les premiers utilisateurs et apprendre d'eux
- Préparer un plan d'analyse
- Soutenir les processus de gestion des données

Cette étude de cas a été préparée par Harpa Isfeld-Kiely du CCNMI et élaborée en collaboration avec les membres de l'équipe Discovery du Centre de gestion, d'innovation et d'analyse des données de l'Agence de la santé publique du Canada, avec des documents fournis par BlueDot Inc.

Numéro de projet CCNMI : 723

ISBN: 978-1-927988-79-4

La publication de ce document a été rendue possible grâce à une contribution financière de l'Agence de la santé publique du Canada aux Centres de collaboration nationale en santé publique (CCNSP). Les opinions exprimées dans ce document ne représentent pas nécessairement celles de l'Agence.