

Cohorte santé et environnement du recensement canadien : guide de l'utilisateur, 2021, version 1b

Date de diffusion : le 18 juin 2025



Statistique
Canada

Statistics
Canada

Canada

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à infostats@statcan.gc.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros suivants :

- | | |
|---|----------------|
| • Service de renseignements statistiques | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur | 1-514-283-9350 |

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « Contactez-nous » > « [Normes de service à la clientèle](#) ».

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, les entreprises, les administrations et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Publication autorisée par la ministre responsable de Statistique Canada

© Sa Majesté le Roi du chef du Canada, représenté par la ministre de l'Industrie, 2025

L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

This publication is also available in English.

Table des matières

1	Contexte	4
2	Données	4
2.1	Recensement de la population, 2021	4
2.2	Base canadienne de données de l'état civil – Décès	5
2.3	Base de données sur les congés des patients.....	6
2.4	Système national d'information sur les soins ambulatoires	6
2.5	Système d'information ontarien sur la santé mentale	7
2.6	Codes postaux annuels pour les adresses postales	7
3	Couplage d'enregistrements dans l'Environnement de couplage de données sociales.....	7
3.1	Méthode de couplage des enregistrements, résultats et qualité	8
3.2	Informations complémentaires	9
4	Dérivation de la Cohorte santé et environnement du recensement canadien de 2021	9
5	Intégration des données	10
5.1	Structures et présentations des fichiers de données	10
5.2	Poids de la cohorte et les poids de répliques	14
5.3	Limites des fichiers analytiques.....	14
6	Accès aux données et Lignes directrices pour la diffusion des données.....	15
6.1	Coordonnées de l'administrateur	15
6.2	Lignes directrices pour la diffusion des données	15
Annexe A – Répartition en pourcentage de certaines caractéristiques démographiques, sociales et économiques de la Cohorte santé et environnement du recensement canadien de 2021 et du fichier du Centre de données de recherche du Recensement de 2021		17
Annexe B – Schéma de la base de données de la Cohorte santé et environnement du recensement canadien de 2021 pour les utilisateurs des données des Centres de données de recherche		18

Cohorte santé et environnement du recensement canadien : guide de l'utilisateur, 2021, version 1b

1 Contexte

Les [Cohortes santé et environnement du recensement canadien](#) (numéro d'enregistrement [5422](#)) sont des fondés sur la population, dans lesquels les données de l'échantillon du questionnaire détaillé du Recensement de la population sont couplées avec les données administratives, y compris celles de la Base canadienne de données de l'état civil – Décès, de la Base de données sur les congés des patients, du Système national d'information sur les soins ambulatoires, du Système d'information ontarien sur la santé mentale et des codes postaux annuels pour les adresses postales. Ces données peuvent être utilisées pour examiner les résultats en matière de santé selon certaines caractéristiques démographiques, sociales et économiques, comme l'âge, le sexe à la naissance et le genre, la scolarité, les régions géographiques, le logement, l'immigration et la diversité ethnoculturelle, et le revenu. Les couplages de microdonnées ([019-2019](#)) ont été approuvés conformément à la [Directive sur le couplage de microdonnées de Statistique Canada](#).

La Cohorte santé et environnement du recensement canadien de 2021 est le plus récent d'une série de couplages de microdonnées basées sur la population. La cohorte se compose de 8 660 377 personnes sélectionnées dans l'échantillon long du recensement de la population de 2021, avec des suivis relatifs à la mortalité, aux hospitalisations, aux soins ambulatoires et aux services de santé mentale. La Cohorte santé et environnement du recensement canadien de 2021 comprend également les codes postaux annuels des adresses postales depuis 1981 pour les membres de la cohorte. Les données couplées peuvent être utilisées pour examiner la mortalité en fonction de caractéristiques démographiques, sociales et économiques telles que l'âge, le sexe à la naissance et le sexe, l'éducation, les zones géographiques, le logement, l'immigration et la diversité ethnoculturelle, le revenu, etc. Les périodes de suivi pour la version 1b de la Cohorte santé et environnement du recensement canadien de 2021 ont été les suivantes :

- Base canadienne de données de l'état civil – Décès : du 11 mai 2021 au 31 décembre 2023
- Base de données sur les congés des patients : du 1^{er} avril 2000 au 31 mars 2024
- Système national d'information sur les soins ambulatoires : du 1^{er} avril 2002 au 31 mars 2024
- Système d'information ontarien sur la santé mentale : du 1^{er} avril 2006 au 31 mars 2024
- Codes postaux annuels pour les adresses postales : de 1981 à 2022.

Le guide de l'utilisateur décrit la Cohorte santé et environnement du recensement canadien de 2021, version 1b.

2 Données

2.1 Recensement de la population, 2021

Dans le cadre du Programme du [Recensement de 2021](#), les ménages canadiens sont dénombrés à l'aide de deux principaux types de questionnaires : le questionnaire abrégé et le questionnaire détaillé. Un échantillon de 25 % des ménages canadiens ont reçu un questionnaire détaillé en 2021, qui inclut les questions du questionnaire abrégé. Les autres ménages ont reçu le questionnaire abrégé. Pour le Recensement de 2021, la date de référence a été fixée au 11 mai. Les renseignements fournis dans les questionnaires du recensement devraient correspondre à la situation de chaque personne au 11 mai 2021, à moins d'avis contraire. Cette date de référence fait en sorte que les renseignements recueillis dans le questionnaire peignent un portrait fidèle de la société canadienne à ce moment précis de notre histoire.

Outre les questions du questionnaire abrégé, le questionnaire détaillé comprenait une série de questions visant à peindre un portrait complet de la population et des ménages canadiens, en fonction de leurs caractéristiques démographiques, sociales et économiques. Les sujets abordés sont les suivants :

- Éducation, formation et apprentissage
- Estimations et projections démographiques
- Familles, ménages et logement
- Immigration et diversité ethnoculturelle
- Industries
- Langues
- Navettage vers le travail
- Peuples autochtones
- Population et démographie
- Professions
- Religion
- Revenu, pensions, dépenses et richesse
- Société et communauté
- Travail.

À partir de 2021, le recensement posera des questions sur le sexe à la naissance et le genre des personnes. Bien que les données sur le sexe à la naissance soient nécessaires pour mesurer certains indicateurs, à partir du recensement de 2021, le genre (et non le sexe) est la variable standard utilisée dans les concepts et les classifications. Pour plus de détails sur le nouveau concept de genre, consultez le [Guide de référence sur l'âge, le sexe à la naissance et le genre, Recensement de la population, 2021](#).

Le [tableau 9.1](#) du Guide du recensement de la population, 2021, Statistique Canada – N° 98-304-X au catalogue montre les taux de réponse obtenus par le traitement et l'évaluation de la qualité des données. À l'échelle nationale, le taux de réponse total était de 95,7 % pour le questionnaire détaillé – logements privés occupés (pondéré). En 2021, 63 subdivisions de recensement définies comme des réserves et des agglomérations ont été partiellement dénombrées. Pour ces réserves et établissements, le dénombrement des logements n'a pas été permis ou n'a pu être complété pour les diverses raisons décrites à l'[annexe 1.5](#) du Guide du recensement de la population, 2021, Statistique Canada – No 98-304-X au catalogue. Pour ces domaines, les comparaisons entre les recensements de 2016 et de 2021 peuvent être moins précises en raison de ces données manquantes.

Pour déterminer le nombre de personnes qui ont été omises ou comptées plus d'une fois, Statistique Canada effectue des études postcensitaires de la couverture de la population recensée, à l'aide d'échantillons représentatifs de la population. Les résultats de ces études sont habituellement disponibles deux ans après le jour du Recensement. Pour le Recensement de 2021, le [taux de sous-dénombrement net](#) (pourcentage de personnes omises moins les personnes recensées plus d'une fois) était de 3,1 %. Ce taux est légèrement supérieur à celui du [Recensement de 2016](#), qui était de 2,4 %. Le [rapport technique sur la couverture](#), Recensement de la population, 2021, Statistique Canada – No 98-304-X au catalogue, porte sur les erreurs de couverture du Recensement de 2021.

Pour le Recensement de 2021, la stratégie de diffusion des indicateurs de qualité a été entièrement revue, dans le but de fournir des informations plus détaillées sur la qualité des données. Pour plus d'informations à ce sujet, les chercheurs sont invités à consulter le document [Lignes directrices sur la qualité des données du Recensement de 2021](#).

La Cohorte santé et environnement du recensement canadien (CSERCan) de 2021 est un sous-ensemble de l'échantillon du questionnaire détaillé du Recensement de 2021 qui ne comprenait que des ménages privés, y compris ceux vivant dans des logements privés rattachés à des logements collectifs au Canada, mais à l'exclusion de ceux qui vivent dans des réserves et des établissements incomplètement dénombrés. Cela signifie que les ménages vivant dans des logements collectifs, à l'extérieur du Canada et dans les 63 réserves et établissements partiellement dénombrés ont été considérés comme « hors champ » pour la CSERCan de 2021.

2.2 Base canadienne de données de l'état civil – Décès

La [Base canadienne de données de l'état civil – Décès \(BCDECD\)](#) est une enquête administrative conçue pour recueillir mensuellement et annuellement, auprès de tous les registres de l'état civil provinciaux et territoriaux, des renseignements démographiques et médicaux (cause de décès) relatifs à tous les décès survenus au Canada. Au moment de la production de la version 1b de la CSERCan de 2021, les données sur les décès de 2021 à 2023 étaient considérées comme provisoires, car des améliorations à la méthodologie et à l'actualité des données ont réduit la durée de la collecte des données. Les données sur les décès de 2021 à 2022 pour le Yukon n'étaient pas disponibles aux fins du couplage des enregistrements au moment du couplage avec le Dépôt d'enregistrements dérivés (DED).

La variable de la cause de décès dans la base de données est classée selon la [Classification statistique internationale des maladies et des problèmes de santé connexes](#) (CIM) de l'Organisation mondiale de la Santé¹. Les données sur les causes de décès de 2021 à 2023 sont codées selon la dixième révision de la CIM (CIM-10).

2.3 Base de données sur les congés des patients

La [Base de données sur les congés des patients \(BDCP\)](#)² contient les données administratives, cliniques et démographiques sur les congés des hôpitaux (y compris les décès survenus dans les hôpitaux, les départs volontaires et les transferts) pour toutes les provinces et tous les territoires, à l'exception du Québec. Au fil du temps, la BDCP a été utilisée pour saisir des données sur les soins de courte durée, y compris les chirurgies d'un jour, les soins de santé chroniques et les soins de réadaptation. Les données de la BDCP pour les exercices 2000-2001 à 2023-2024 ont été incluses dans le couplage (exercices comprenant la période du 1^{er} avril 2000 au 31 mars 2024). La BDCP est un fichier de données fondé sur les événements, ce qui signifie qu'il peut y avoir plus d'un enregistrement par personne.

Dans la BDCP, les instructions relatives à la collecte des éléments d'information peuvent varier d'un secteur de compétence à l'autre et évoluer au fil du temps; notamment, la collecte de certains éléments d'information peut être obligatoire, obligatoire si applicable, facultative ou non applicable, selon le secteur de compétence. Pour en savoir davantage sur les éléments d'information et la couverture de la BDCP, veuillez consulter la section « Éléments de données » du [site Web des métadonnées de la BDCP](#). Pour obtenir plus de renseignements sur les méthodologies utilisées pour la BDCP, consultez le [Système global de classification ambulatoire \(SGCA\)](#), le [Guide de soutien à la prise de décisions pour les groupes clients : GMA+](#) et la [Norme relative au contenu de données sur les soins ambulatoires et de courte durée](#).

2.4 Système national d'information sur les soins ambulatoires

Le [Système national d'information sur les soins ambulatoires \(SNISA\)](#)³ contient des données sur les soins ambulatoires donnés en milieu hospitalier et au sein de la collectivité, y compris pour les chirurgies d'un jour, les cliniques communautaires et les cliniques de consultation externe, et les services d'urgence. La collecte des données sur les visites de patients se fait au moment de la prestation du service dans les établissements participants de plusieurs secteurs de compétence. Les données du SNISA pour les exercices 2002-2003 à 2023-2024 sont incluses dans le couplage, y compris pour la période du 1^{er} avril 2002 au 31 mars 2024. Le SNISA est un fichier de données fondé sur les événements, ce qui signifie qu'il peut y avoir plus d'un enregistrement par personne.

La méthodologie de regroupement utilisée pour le SNISA est celle du Système global de classification ambulatoire (SGCA), un système national de regroupement des données des patients en soins ambulatoires qui ont soumises à la BDCP ou au SNISA. Les regroupements comprennent les visites au service d'urgence, les interventions en soins ambulatoires, les services de réadaptation et les visites en clinique, ainsi que des cas de regroupement exceptionnel, comme les consultations téléphoniques et l'imagerie diagnostique directe. Les patients sont regroupés selon un certain nombre d'éléments d'information, y compris le diagnostic ou l'intervention (pour la BDCP), et l'indicateur de visite au service d'urgence, le mode de consultation, l'issue de la visite, le type de soins ambulatoires ou le domaine de programme (pour le SNISA).

L'Institut canadien d'information sur la santé n'est pas tenu de soumettre des données au SNISA, et les instructions relatives à la collecte des éléments d'information peuvent varier d'un secteur à l'autre et changer au fil du temps. Comme pour la BDCP, la collecte des différents éléments d'information peut être obligatoire, obligatoire si applicable, facultative ou non applicable, selon le secteur de compétence. Pour en savoir plus sur la couverture des données provinciales et la collecte des éléments d'information, veuillez consulter la [Couverture des données](#) et la section « Étendue des données » du [site Web des métadonnées du SNISA](#). Pour obtenir plus de renseignements sur les méthodologies utilisées pour le SNISA, consultez le [Système global de classification ambulatoire \(SGCA\)](#) et la [Norme relative au contenu de données sur les soins ambulatoires et de courte durée](#).

1. Organisation mondiale de la Santé. (s. d.). [Classification statistique internationale des maladies et des problèmes de santé connexes \(CIM\)](#)
 2. Institut canadien d'information sur la santé. (s. d.). [Métadonnées de la Base de données sur les congés des patients \(BDCP\)](#)
 3. Institut canadien d'information sur la santé. (s. d.). [Métadonnées du Système national d'information sur les soins ambulatoires \(SNISA\)](#)

2.5 Système d'information ontarien sur la santé mentale

Le [Système d'information ontarien sur la santé mentale \(SIOSM\)](#)⁴ contient des données de dossiers de patients pour toutes les personnes recevant des services de soins de santé mentale pour adultes en Ontario, de même que pour certaines personnes recevant des services dans des établissements disposant de lits réservés aux jeunes patients hospitalisés et dans certains établissements d'autres provinces, à compter de l'exercice 2006-2007. Les données du SIOSM sont recueillies à l'aide de la version 2.0 de l'instrument d'évaluation des résidents – santé mentale (RAI-MH®), et comprennent des renseignements sur l'admission des patients ainsi que l'évaluation, la planification des soins et la mesure des résultats. Pour obtenir de plus amples renseignements sur les méthodologies relatives aux SIOSM, comme les méthodologies de la composition des cas, consultez le [Système de classification des patients hospitalisés en psychiatrie \(SCIPP\)](#) élaboré pour le RAI-MH. Le SIOSM est un fichier de données fondé sur les événements, ce qui signifie qu'il peut y avoir plus d'un enregistrement par personne.

Pour ce couplage d'enregistrements, les enregistrements du SIOSM couvrant les exercices de 2006-2007 à 2023-2024 ont été inclus. Les données sont actuellement soumises à l'Institut canadien d'information sur la santé par les hôpitaux participants de l'Ontario, ainsi que par trois établissements de Terre-Neuve-et-Labrador et un établissement du Manitoba.

Pour en savoir davantage sur les éléments d'information et la qualité des données du SIOSM, veuillez consulter les sections « Éléments de données » et « Qualité des données » du [site Web de métadonnées du SIOSM](#).

2.6 Codes postaux annuels pour les adresses postales

Au Canada, les déclarations de revenus sont soumises annuellement à l'Agence du revenu du Canada (ARC). Le Fichier maître des particuliers T1 (FMPT1), aussi connu sous le nom de T1 Général et Annexes, est une collection des déclarations de revenus que l'ARC partage avec Statistique Canada, et il fournit des renseignements sur le revenu et les caractéristiques démographiques (p. ex. date de décès) sur les déclarants au Canada. Chaque résident du Canada qui gagne un revenu imposable est tenu de produire une déclaration de revenus, appelée formulaire T1, à la fin de l'année au cours de laquelle le revenu a été reçu. Par conséquent, le FMPT1 comprend presque toutes les personnes qui ont produit une déclaration de revenus T1 pour l'année de référence (c.-à-d. certains déclarants tardifs pourraient ne pas être inclus) ou qui ont reçu des prestations fiscales canadiennes pour enfants (PFCE) et leurs conjoints non déclarants.

Le FMPT1 (de 1981 à 2022) est la principale source de données pour les codes postaux annuels pour les adresses postales. Les codes postaux de l'adresse postale déclarés sur ces dossiers fiscaux ont été extraits pour estimer le lieu de résidence d'une personne pour cette année de référence. Veuillez noter que pour certains déclarants, les adresses postales utilisées pour remplir les documents fiscaux T1 peuvent ne pas être associées à leur lieu de résidence (p. ex. cases postales, bureaux de comptables ou d'avocats, adresses des parents pour les jeunes adultes, adresses des enfants pour les parents âgés).⁵

3 Couplage d'enregistrements dans l'Environnement de couplage de données sociales

L'[Environnement de couplage des données sociales \(ECDS\)](#) a produit une série de fichiers de population couplés pour l'analyse sociale en utilisant le [Dépôt d'enregistrements dérivés \(DED\)](#), une base de données relationnelle dynamique qui ne contient que des identifiants personnels de base. Les données d'enquête et les données administratives sont reliées au DED à l'aide du logiciel SAS et du logiciel [G-Coup \(Couplage d'enregistrements - Système Généralisé\)](#), un logiciel de couplage d'enregistrements généralisé basé sur SAS qui prend en charge le couplage déterministe et probabiliste développé par Statistique Canada. Voir le site Web de Statistique Canada pour plus d'informations concernant l'ECDS.

La Section 3.1 du guide de l'utilisateur présente les informations pertinentes extraites des rapports méthodologiques ECDS. La Section 3.2 fournit des informations supplémentaires sur les données qui ne sont pas disponibles dans ces rapports.

4. Institut canadien d'information sur la santé. (s. d.). [Métadonnées du Système d'information ontarien sur la santé mentale \(SIOSM\)](#)

5. [Comparaison du lieu de résidence entre le fichier T1 sur la famille et le recensement : Évaluation à partir d'un couplage d'enregistrements](#)

3.1 Méthode de couplage des enregistrements, résultats et qualité

En bref, 96,1 %⁶ des enregistrements de personnes dans le fichier de la Base de données des réponses (BDR) du Recensement de 2021 ont été couplés de manière probabiliste aux enregistrements de personnes dans le DED. Il correspond au taux de couplage des données du formulaire abrégé du recensement. Le taux de couplage des données de l'échantillon du formulaire détaillé du recensement était supérieur à celui des données du formulaire abrégé de 0,6 point de pourcentage, soit 96,7 %. Le couplage probabiliste des enregistrements s'est appuyé sur des informations telles que le nom, le prénom, le sexe à la naissance, la date de naissance, le numéro de téléphone, et la géographie (aussi détaillée que le code postal à 6 caractères). Le pourcentage global estimé de fausses correspondances était de 0,39 % sur la base des données du formulaire abrégé du recensement.

Les enregistrements de décès figurant dans La Base canadienne de données de l'état civil – Décès (BCDECD) de 2021 à 2023 ont été reliés aux enregistrements de personnes dans le DED au moyen d'une méthode probabiliste de couplage d'enregistrements utilisant des informations telles que le nom de famille, le prénom, le sexe à la naissance, la date de naissance, la date de décès et la géographie (aussi détaillée que le code postal à 6 caractères). Environ 99,1 % des enregistrements de décès ont été couplés aux enregistrements de personnes dans le DED. Le pourcentage global estimé de fausses correspondances devrait être inférieur à 0,3 %.

Pour les données fondées sur des événements, comme celles de la Base de données sur les congés des patients (BDCP), du Système national d'information sur les soins ambulatoires (SNISA) et du Système d'information ontarien sur la santé mentale (SIOSM), une méthodologie de couplage d'enregistrements déterministe en deux étapes est utilisée afin d'apparier les enregistrements de chacun des fichiers de données avec les enregistrements des personnes dans le Dépôt d'enregistrements dérivés (DED), à l'aide d'un ou de plusieurs identificateurs comme le sexe à la naissance, la date de naissance, et le code postal, et/ou le numéro d'assurance maladie/carte de santé et la province ou le territoire de délivrance. Chaque événement (congé de l'hôpital, visite ambulatoire et/ou l'évaluation de santé mentale) a été couplé de façon déterministe au DED pour tenir compte de variables de couplage pouvant changer au fil du temps, comme le code postal de la résidence, et l'information sur la carte d'assurance-maladie.

Les taux de couplage cumulés pour la BDCP, le SNISA et le SIOSM sont les suivants :

- BDCP : taux de couplage cumulatif d'environ 95,2 % des dossiers de congés des patients hospitalisés de 2000-2001 à 2023-2024
- SNISA : taux de couplage cumulatif d'environ 95,6 % des dossiers de soins ambulatoires de 2002-2003 à 2023-2024
- SIOSM : taux de couplage cumulatif de 94,6 % des dossiers de services de santé mentale de 2006-2007 à 2023-2024.

Les pourcentages de fausses correspondances ne sont pas estimés pour ces liens d'enregistrement déterministes.

Rappelons que les adresses postales utilisées pour enregistrer les déclarations de revenus T1 constituent la principale source de données pour l'historique des codes postaux. Les enregistrements de personnes du le Fichier maître des particuliers T1 (FMPT1) ont été reliés aux enregistrements de personnes du DED à l'aide d'une méthode probabiliste basée sur le numéro d'assurance sociale, le nom de famille, le prénom, le sexe, la date de naissance, la date de décès, le numéro de téléphone et la géographie (aussi détaillée que le code postal à 6 caractères). Environ 99,7 % des 46 millions de personnes enregistrées dans le FMPT1 (de 1981 à 2022) ont été soit couplées aux enregistrements de personnes dans le DED, soit ajoutées au DED en tant que nouvelles personnes. Les pourcentages estimés de fausses correspondances au cours des années fiscales n'ont jamais été inférieurs à 0,21 %.

6. Les enregistrements de personnes en double sont exclus du calcul.

3.2 Informations complémentaires

Cette section fournit des informations supplémentaires sur les données de l'échantillon complet du recensement de 2021 nécessaires pour produire la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021. Ces données ont été extraites des fichiers suivants :

- Le fichier de la Base de données des réponses (BDR) du Recensement de 2021
- Le fichier du Centre de données de recherche (CDR) du Recensement de 2021,

cette dernière est la version diffusée des données du formulaire long du recensement. Environ 1,17 % des enregistrements de personnes dans le fichier du CDR du Recensement de 2021 n'existaient pas dans la version du fichier de la BDR du Recensement de 2021 ayant fait l'objet des couplages décrits dans les rapports méthodologiques. Plus de 90 % d'entre elles étaient des personnes qui avaient rempli le questionnaire détaillé (formulaire [2A-R](#)) envoyé aux ménages privés dans les communautés des Premières Nations, les établissements métis, les régions inuites et d'autres régions éloignées afin de recenser 100 % de la population. Comme ces enregistrements de personnes n'ont jamais été liés aux enregistrements de personnes dans le DED, ils ont été considérés comme « hors champ » pour la CSERCan de 2021.

4 Dérivation de la Cohorte santé et environnement du recensement canadien de 2021

La Cohorte santé et environnement du recensement canadien (CSERCan) de 2021 a été créée à partir du fichier du Recensement de 2021 des centres de données de recherche (CDR) et des clés de couplage entre le fichier de la Base de données des réponses (BDR) du Recensement de 2021 et des fichiers de la Base canadienne de données sur l'état civil – Décès (BCDECD) de 2021 à 2023 produite à l'aide du Dépôt d'enregistrements dérivés (DED). Cette dernière cohorte, composée de 8 660 377 répondants, comprenait 95,45 % de tous les enregistrements personnes figurant dans le fichier du Recensement de 2021 des CDR. Il s'agit d'une cohorte fixe, ce qui signifie que les membres de la cohorte sont définis lors de la création de la cohorte et qu'aucun nouveau membre n'est ajouté pendant la période de suivi. Les étapes utilisées pour dériver la CSERCan de 2021 sont présentées dans le Tableau 1.

Tableau 1
Dérivation de la Cohorte santé et environnement du recensement canadien de 2021

Description	Taille ¹
Cohorte de base	Nombre d'enregistrements de personnes
Enregistrements de personnes dans le fichier du Centre de données de recherche (CDR) du Recensement de 2021	9 072 940
Critères d'exclusion (dans l'ordre des suppressions de la cohorte de base)	Pourcentage de la taille de la cohorte de base
1. Enregistrements de personnes ne figurant pas dans la version du fichier de la Base de données des réponses (BDR) du Recensement de 2021 qui a été liée au Dépôt d'enregistrements dérivés (DED)	-1,17
2. Enregistrements de personnes non liés au DED	-3,06
3. Enregistrements de personnes présumés en double selon le couplage interne d'enregistrements du fichier de la BDR au Recensement de 2021	-0,30
4. Enregistrements de personnes liés à un décès survenu avant le jour du Recensement (mai 11 2021)	-0,01
5. Enregistrement de personne présumé faussement couplé au DED parmi ceux liés aux décès survenus à partir du jour du recensement	-0,01
Cohorte finale	Number of person records
Enregistrements des personnes de la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021	8 660 377

1. Recensé par l'identificateur de recensement pour le tableau des personnes (PP_ID).

Sources : Statistique Canada, Cohorte santé et environnement du recensement canadien de 2021 (numéro d'enregistrement [5422](#)), Recensement de la population, 2021 (numéro d'enregistrement [3901](#)) et Base canadienne de données de l'état civil – Décès, de 2021 à 2023 (numéro d'enregistrement [3233](#)).

5 Intégration des données

Les fichiers de la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021 conçus pour être fusionnés entre eux ainsi qu'avec le fichier du Centre de données de recherche (CDR) du Recensement de 2021 et les fichiers des résultats en matière de santé, afin de créer des fichiers à des fins de recherche et d'analyse.

Étant donné le nombre important d'enregistrements et de variables dans ces fichiers – par exemple, le fichier du CDR du recensement de 2021 contient plus de 9 millions d'enregistrements et 907 variables - il est recommandé aux chercheurs de ne conserver que les variables du fichier qui sont nécessaires pour répondre à leurs questions de recherche spécifiques.

Il est à noter que certains ensembles de données sources peuvent avoir les mêmes concepts en matière de variables d'une année à l'autre, mais utiliser des appellations, des formats et des codages différents (p. ex. genre, âge, date de naissance). Veuillez noter que les ensembles de données pourraient contenir les mêmes noms de variable tout en utilisant des formats différents. Consultez les dictionnaires de données et renommez et/ou supprimez les variables inutiles avant de fusionner les fichiers.

La variable **UniqID** contient des numéros d'identification uniques pour chacun des membres de la cohorte. Notez que ces numéros d'identification sont uniques au sein de cette cohorte.

Pour les utilisateurs des données du CDR, un schéma de la fusion des fichiers de la CSERCan de 2021 est présenté à la Figure 1, Annexe B. Si vous avez des questions ou des inquiétudes concernant la fusion des fichiers de données, veuillez contacter l'analyste de votre CDR.

5.1 Structures et présentations des fichiers de données

Le Tableau 2 présente les derniers fichiers de la CSERCan de 2021 disponibles dans les CDR.

Tableau 2

Liste des derniers fichiers de la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021 disponibles dans les Centres de données de recherche

Nom du fichier de données	Nombre d'enregistrements	Description
canhec_c21keyfile_f3_v1b	8 660 377	Fichier des clés de couplage entre le fichier du Centre de données de recherche (CDR) du Recensement de 2021 et les fichiers de la Base canadienne de données de l'état civil – Décès (BCDECD) de 2021 à 2023, constitué des personnes incluses dans la CSERCan de 2021.
canhec_c21ltfufile_f3_v1b	8 660 377	Fichier des enregistrements perdus aux fins de suivi selon les clés de couplage entre le fichier du CDR du recensement de 2021 et le Dépôt d'enregistrements dérivés (DED), extraites le 23 janvier 2025.
canhec_c21weights_f3_v1b	8 660 377	Fichier contenant les poids de cohorte et de rééchantillonnage pour la CSERCan de 2021
canhec_c21mobility_f3_v1	8 624 545	Fichier contenant les codes postaux annuels des adresses postales (de 1981 à 2022) pour la CSERCan de 2021
canhec_c21dad00_f3_v1–canhec_c21dad23_f3_v1	12 644 011	Fichiers analytiques (masqués) de la Base de données sur les congés des patients (BDGP), par exercice financier, pour la CSERCan de 2021
canhec_c21dadinstcode00_f3_v1–canhec_c21dadinstcode23_f3_v1	1 264 011	Numéros d'établissement attribués aux enregistrements de la BDGP dans les fichiers analytiques (voir la Section 6.2)
canhec_dadinstcode0023_f3_v1	1 472	Liste des établissements participant à la BDGP dans les fichiers analytiques (voir la Section 6.2)
canhec_c21nacrs02_f3_v1–canhec_c21nacrs23_f3_v1	58 123 412	Fichiers analytiques (masqués) du Système national d'information sur les soins ambulatoires (SNISA), par exercice financier, pour la CSERCan de 2021
canhec_c21nacrsinstcode00_f3_v1–canhec_c21nacrsinstcode23_f3_v1	58 123 412	Numéros d'établissement attribués aux enregistrements du SNISA dans les fichiers analytiques (voir la Section 6.2)
canhec_nacrsinstcode0223_f3_v1	618	Liste des établissements participant au SNISA dans les fichiers analytiques (voir la Section 6.2)
canhec_c21omhrs0623_f3_v1	277 529	Fichier analytique (masqué) du Système d'information ontarien sur la santé mentale (SIOSM) pour la CSERCan de 2021
canhec_omhrsinstcode0623_f3_v1	109	Liste des établissements participant au SIOSM dans le fichier analytique (voir la Section 6.2)

Sources : Statistique Canada, Cohorte santé et environnement du recensement canadien de 2021 (numéro d'enregistrement [5422](#)), Recensement de la population, 2021 (numéro d'enregistrement [3901](#)) et Base canadienne de données de l'état civil – Décès, de 2021 à 2023 (numéro d'enregistrement [3233](#)); Agence du revenu du Canada, Fichier maître des particuliers T1; et Institut canadien d'information sur la santé, Base de données sur les congés des patients, de 2000–2001 à 2023–2024 ([métadonnées](#)), Système national d'information sur les soins ambulatoires, de 2002–2003 à 2023–2024 ([métadonnées](#)) et Système d'information ontarien sur la santé mentale, de 2006–2007 à 2023–2024 ([métadonnées](#)).

Le fichier **canhec_c21keyfile_f3_v1b** est le fichier des clés de couplage entre le fichier du Recensement de 2021 des CDR et les fichiers de la Base canadienne de données sur l'état civil – Décès de 2021 à 2023. Ce fichier contient 8 660 377 enregistrements; ce qui correspond à la taille de la cohorte. Il contient également des identificateurs d'enregistrement pour 150 065 décès. Veuillez noter que les identificateurs **UniqID** et **PP_ID** ont une relation univoque. De plus, **EVENT_YEAR**, **PLACEOFDEATH_PROVINCE** et **REGISTRATION_NUMBER** identifient ensemble de façon unique un enregistrement de BCDECD. Le Tableau 3 décrit les variables du fichier.

Tableau 3

Liste des variables et leurs descriptions pour le fichier des clés de couplage entre le fichier du Centre de données de recherche du Recensement de 2021 et les fichiers de la Base canadienne de données de l'état civil – Décès de 2021 à 2023, constitué des personnes incluses dans la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021

Variable	Type	Longueur	Description
UniqID	Num.	8	Identificateur unique CSERCan
PP_ID	Num.	8	Identifiant pour la table personne du recensement
PRCDDA	Num.	8	Province, Division du recensement, Secteur de la diffusion
HhNum	Num.	8	Clé pour le ménage
PpNum	Num.	8	Clé pour les personnes
EVENT_YEAR	Carac.	4	Année où le décès est survenu (AAAA)
PLACEOFDEATH_PROVINCE	Carac.	3	Province/territoire où le décès est survenu
REGISTRATION_NUMBER	Carac.	6	Numéro d'enregistrement du décès

Sources : Statistique Canada, Cohorte santé et environnement du recensement canadien de 2021 (numéro d'enregistrement [5422](#)), Recensement de la population, 2021 (numéro d'enregistrement [3901](#)) et Base canadienne de données de l'état civil – Décès, de 2021 à 2023 (numéro d'enregistrement [3233](#)).

Le fichier **canchec_c21ltfufile_f3_v1b** est le fichier qui contient la cohorte complète (8 660 377 enregistrements) avec un indicateur qui identifie les personnes qui ne sont plus couplés au Dépôt d'enregistrements dérivés (DED) en raison du nettoyage des données et des révisions apportées au dépôt au fil du temps. Ces membres de la cohorte sont donc considérés comme « perdus aux fins de suivi ». Le pourcentage de personnes perdues aux fins de suivi en date de la version 1b de la CSERCan de 2021 est inférieur à 0,01 %⁷. Le tableau 4 décrit les variables du fichier. Le Tableau 4 décrit les variables du fichier.

Tableau 4

Liste des variables et leurs descriptions pour le fichier des enregistrements perdus aux fins de suivi pour la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021, selon les clés de couplage entre le fichier du Centre de données de recherche du recensement de 2021 et le Dépôt d'enregistrements dérivés, extraites le 23 janvier 2025

Variable	Type	Length	Description
UniqID	Num.	8	Identificateur unique de la CSERCan
LTFU	Num.	8	Indicateur pour signaler les enregistrements perdus aux fins de suivi, où 1 = perdu au moment du suivi et 0 = non perdu au moment du suivi

Sources : Statistique Canada, Cohorte santé et environnement du recensement canadien de 2021 (numéro d'enregistrement [5422](#)), Recensement de la population, 2021 (numéro d'enregistrement [3901](#)) et Base canadienne de données de l'état civil – Décès, de 2021 à 2023 (numéro d'enregistrement [3233](#)).

Le fichier **canchec_c21weights_f3_v1b** est celui qui contient la cohorte et les poids de rééchantillonnage pour la CSERCan de 2021. Il contient également 8 660 377 personnes; ce qui correspond à la taille de la cohorte. Le Tableau 5 décrit les variables du fichier. Pour obtenir de plus amples renseignements sur ces facteurs de pondération, veuillez consulter la Section 5.2.

Tableau 5

Liste des variables et leurs descriptions pour le fichier contenant les poids de cohorte et de rééchantillonnage pour la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021

Variable	Type	Longueur	Description
UniqID	Num.	8	Identificateur unique CSERCan
CanCHECW2	Num.	8	Poids de la cohorte
CanCHEC_REPWT1–CanCHEC_REPWT100	Num.	8	100 poids de répliques

Sources : Statistique Canada, Cohorte santé et environnement du recensement canadien de 2021 (numéro d'enregistrement [5422](#)), Recensement de la population, 2021 (numéro d'enregistrement [3901](#)) et Base canadienne de données de l'état civil – Décès, de 2021 à 2023 (numéro d'enregistrement [3233](#)).

7. Plus de la moitié des enregistrements perdus aux fins de suivi concernaient des femmes+ (58,0 %), tandis que 45,4 % de ces enregistrements étaient en Ontario le jour du Recensement et presque tous étaient âgés de plus de 15 ans sur cette journée.

Le fichier **canchec_c21mobility_f3_v1** contient les codes postaux annuels des adresses postales de 1981 à 2022 pour la CSERCan de 2021. Il contient 8 624 545 enregistrements. Le Tableau 6 décrit les variables du fichier.

Tableau 6

Liste des variables et leurs descriptions pour le fichier contenant les codes postaux annuels des adresses postales (de 1981 à 2022) pour la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021

Variable	Type	Longueur	Description
UniqID	Num.	8	Identificateur unique de la CSERCan
PC1981-PC2022	Carac.	6	Code postal par année de validité selon le Dépôt d'enregistrements dérivés

Sources : Statistique Canada, Cohorte santé et environnement du recensement canadien de 2021 (numéro d'enregistrement 5422), Recensement de la population, 2021 (numéro d'enregistrement 3901) et Base canadienne de données de l'état civil – Décès, de 2021 à 2023 (numéro d'enregistrement 3233); et Agence du revenu du Canada, Fichier maître des particuliers T1.

Le fichier **canchec_c21dad00_f3_v1-canchec_c21dad23_f3_v1** contient les enregistrements la Base de données sur les congés des patients (BDCP) par exercice pour la CSERCan de 2021. Il contient 12 644 011 enregistrements, mais le nombre annuel d'enregistrements varie selon l'exercice. Le Tableau 7 décrit les variables contenues dans les fichiers.

Tableau 7

Liste des variables et leurs descriptions pour les fichiers analytiques (masqués) de la Base de données sur les congés des patients (BDCP), par exercice, pour la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021

Variable	Type	Longueur	Description
UniqID	Num.	8	Identificateur unique de la CSERCan
DAD_ID	Carac.	15	Identificateur d'enregistrement de la BDCP (randomisé)
FISCAL_YEAR et autres variables de la BDCP	Num./Carac.	Varie	Exercice de l'enregistrement de congé de l'hôpital et autres variables de la BDCP

Sources : Statistique Canada, Cohorte santé et environnement du recensement canadien de 2021 (numéro d'enregistrement 5422), Recensement de la population, 2021 (numéro d'enregistrement 3901) et Base canadienne de données de l'état civil – Décès, de 2021 à 2023 (numéro d'enregistrement 3233); et Institut canadien d'information sur la santé, Base de données sur les congés des patients, de 2000–2001 à 2023–2024 ([métadonnées](#)).

Le fichier **canchec_c21nacrs02_f3_v1-canchec_c21nacrs23_f3_v1** contient les enregistrements du Système national d'information sur les soins ambulatoires (SNISA) par exercice pour la CSERCan de 2021. Il contient 58 123 412 enregistrements, mais le nombre d'enregistrements varie selon l'exercice. Le Tableau 8 décrit les variables contenues dans les fichiers.

Tableau 8

Liste des variables et leurs descriptions pour les fichiers analytiques (masqués) du Système national d'information sur les soins ambulatoires (SNISA), par exercice, pour la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021

Variable	Type	Longueur	Description
UniqID	Num.	8	Identificateur unique de la CSERCan
NACRS_ID	Carac.	15	Identificateur d'enregistrement du SNISA (randomisé)
FISCAL_YEAR et autres variables du SNISA	Num./Carac.	Varie	Exercice du dossier des soins ambulatoires ou d'urgence et autres variables du SNISA

Sources : Statistique Canada, Cohorte santé et environnement du recensement canadien de 2021 (numéro d'enregistrement 5422), Recensement de la population, 2021 (numéro d'enregistrement 3901) et Base canadienne de données de l'état civil – Décès, de 2021 à 2023 (numéro d'enregistrement 3233); et Institut canadien d'information sur la santé, Système national d'information sur les soins ambulatoires, de 2002–2003 à 2023–2024 ([métadonnées](#)).

Le fichier **canchec_c21omhrs0623_f3_v1** contient les enregistrements du Système d'information ontarien sur la santé mentale (SIOSM) pour la CSERCan de 2021. Il contient 277 529 enregistrements. Le Tableau 9 décrit les variables du fichier.

Tableau 9**Liste des variables et leurs descriptions pour Fichier analytique (masqué) du Système d'information ontarien sur la santé mentale (SIOSM) pour la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021**

Variable	Type	Longueur	Description
UniqID	Num.	8	Identificateur unique de la CSERCan
OMHRS_ID	Carac.	15	Identificateur d'enregistrement du SIOSM (randomisé)
FACILITY_PROVINCE et autres variables du SIOSM	Num./Carac.	Varie	Province ou territoire de l'établissement déclarant et autres variables du SIOSM

Sources : Statistique Canada, Cohorte santé et environnement du recensement canadien de 2021 (numéro d'enregistrement [5422](#)), Recensement de la population, 2021 (numéro d'enregistrement [3901](#)) et Base canadienne de données de l'état civil – Décès, de 2021 à 2023 (numéro d'enregistrement [3233](#)); et Institut canadien d'information sur la santé, Système d'information ontarien sur la santé mentale, de 2006–2007 à 2023–2024 ([métadonnées](#)).

5.2 Poids de la cohorte et les poids de répliques

Un fichier des poids (**canchec_c21weights_f3_v1b**) a été créé pour la cohorte afin de la rendre plus représentative de la population cible et de réduire le biais attribuable aux appariements manqués. En bref, le poids principal l'enquête-échantillon du questionnaire détaillé (**CompW2**) a été ajusté à partir des paramètres du modèle en fonction de la probabilité du couplage avec un enregistrement de personne dans le Dépôt d'enregistrements dérivés (DED). Ce poids ajusté a ensuite été calé aux chiffres du recensement afin d'obtenir le poids de la cohorte (**CanCHECW2**).

Rappelons que l'[estimateur par répliques choisi pour l'enquête-échantillon du questionnaire détaillé de 2021 est dérivé de la méthode des demi-échantillons équilibrés de Fay](#). La méthode détermine la création des répliques, le calcul des poids de répliques (100) ainsi que le facteur multiplicatif utilisé pour estimer la variance. L'ensemble de 100 poids de répliqués pour l'enquête-échantillon du questionnaire détaillé (**MAIN_REPWT1–MAIN_REPWT100**) a été ajusté en utilisant la même procédure que ci-dessus pour produire l'ensemble final de 100 poids de répliqués pour la cohorte (**CanCHEC_REPWT1–CanCHEC_REPWT100**).

Pour utiliser ces poids, on doit suivre une procédure d'analyse étayée de l'enquête. La méthode servant à estimer la variance devrait s'appuyer sur l'option des répliques répétées équilibrées (RRÉ) et devrait un ajustement de Fay de 2 ou une valeur epsilon de 0,2928932, selon le logiciel :

- Dans la version 9.3 du SAS ou une procédure ultérieure de SURVEY, il faut indiquer la valeur epsilon comme suit : **VARMETHOD=BRR (FAY=0.2928932)**.
- Dans SUDAAN, utilisez **REPWT CanCHEC_REPWT1–CanCHEC_REPWT100 / ADJFAY=2**.
- Utilisation du paquetage R « [survey : Analysis of Complex Survey Samples](#) », les arguments pour le type de poids de réplification et le facteur de réduction pour les poids dans l'estimateur de variance de Fay dans un plan de RRÉ sont spécifiés comme **type=Fay** et **rho=0.2928932**, respectivement.

Les poids de cohorte et les poids répliques de la CSERCan de 2021 sont disponibles sur **canchec_c21weights_f3_v1b** et le fichier peut être fusionné d'autres fichiers en utilisant la variable **UniqID**.

L'Annexe A présente la répartition en pourcentage de certaines caractéristiques démographiques, sociales et économiques de la CSERCan de 2021 (non pondérée et pondérée) et du fichier du CDR du Recensement de 2021 (pondérée).

5.3 Limites des fichiers analytiques

Le couplage des données s'accompagne d'un biais d'inclusion potentiel. Malgré les poids de cohorte qui ont été conçus pour atténuer ce risque, il pourrait y avoir un biais inconnu si les enregistrements ne faisant pas partie de la cohorte différaient systématiquement de ceux inclus.

La CSERCan de 2021 doit être considérée comme une population en meilleure santé que la population générale en raison de l'exclusion de la population institutionnalisée et vivant dans un logement collectif non institutionnel au niveau de référence.

Lors du couplage des enregistrements, les données de 2021 à 2023 de la BCDECD, ont été considérées comme préliminaire, car des améliorations apportées à la méthodologie et à l'actualité des données ont réduit la durée de la collecte des données. Les données pour le Yukon entre 2021 et 2022 dans la BCDECD n'étaient pas disponibles au moment du couplage d'enregistrements. Les données seront révisées lors des diffusions ultérieures. Malgré le taux de couplage global entre la BCDECD et le DED de 99,1 % pour 2021 à 2023, le taux de couplage est connu pour être plus faible chez les personnes décédées âgées de 7 ans et moins, en particulier celles âgées de moins d'un an, en raison du sous-dénombrement du DED.

Le fichier des codes postaux ne rend pas nécessairement compte de l'endroit où une personne vit, parce que pour certains déclarants, les adresses postales utilisées pour produire les déclarations de revenus T1 peuvent ne pas correspondre à leur lieu de résidence (p. ex. case postale, bureau de comptable ou d'avocat, adresse des parents pour les jeunes adultes, adresse des enfants pour les parents âgés).

Les données contenues dans la BDCP, le SNISA et le SIOSM ne couvrent pas les données de toutes les provinces et de tous les territoires. De plus, la couverture peut varier selon l'exercice et peut ne pas être comparable d'un secteur de compétence à l'autre. Plus précisément, la BDCP ne contient pas de données sur les congés des patients des hôpitaux et les chirurgies d'un jour au Québec.

Les fichiers de données analytiques de la BDCP, du SNISA et du SIOSM contiennent des éléments d'information relatifs aux décès (p. ex. date du décès) et ces renseignements peuvent être incohérents au sein de ces fichiers de données ou entre eux lorsque les événements sont regroupés par patient. De plus, les renseignements sur les décès saisis dans les fichiers de données de la BDCP, du SNISA et du SIOSM peuvent différer des renseignements couplés sur les décès saisis par l'entremise de la BCDEC – Décès. Les renseignements sur les décès tirés de la BCDEC – Décès devraient être considérés comme la norme pour l'analyse des données couplées comportant une mortalité.

Enfin, le statut d'émigration des membres de la cohorte n'est pas connu.

6 Accès aux données et Lignes directrices pour la diffusion des données

6.1 Coordonnées de l'administrateur

Le Centre d'intégration des données de santé et des mesures directes est la division administratrice des Cohortes santé et environnement du recensement canadien (CSERCAN). Les questions concernant les fichiers de données de cohorte peuvent être envoyées au Service clientèle : hd-ds@statcan.gc.ca.

Veuillez-vous reporter au [Programme des centres de données de recherche \(CDR\)](#) pour un accès par l'entremise des CDR.

6.2 Lignes directrices pour la diffusion des données

Pour plus d'informations sur les procédures et les directives concernant le contrôle de la confidentialité, veuillez consulter le dernier document intitulé Lignes directrices pour les Cohortes santé et environnement du Recensement du Canada. Le document a été mis à jour pour la dernière fois en août 2024.

Les données de la CSERCAN ne devraient pas être utilisées pour produire des totalisations du recensement au-delà du « Tableau 1 » descriptif qu'on peut trouver dans un article de journal. Toutes les autres totalisations du recensement doivent être obtenues par l'entremise des données complètes du recensement.

Les données couplées ne doivent pas être utilisées pour produire des statistiques officielles sur la mortalité. Les chiffres et les taux officiels concernant la mortalité sont présentés sur le site Web de Statistique Canada; ils peuvent également être générés à l'aide de la Base canadienne de données sur l'état civil – Décès, accessible auprès des CDR ou sur présentation d'une demande de tableau personnalisé à Statistique Canada (hd-ds@statcan.gc.ca). De plus, les données couplées ne doivent pas être utilisées pour produire des statistiques officielles sur les hospitalisations (soins de courte durée, soins ambulatoires, services de santé mentale). Les nombres et taux officiels d'hospitalisations sont disponibles auprès de l'Institut canadien d'information sur la santé.

REMARQUE : Toutes les analyses et tous les tableaux produits à partir des données de décès de 2021 à 2022 doivent s'accompagner de la note suivante : « Lors du couplage des enregistrements, les données de 2021 à 2023 de la Base canadienne de données de l'état civil – Décès (BCDECD), ont été considérées comme préliminaire, car des améliorations apportées à la méthodologie et à l'actualité des données ont réduit la durée de la collecte des données. Les données pour le Yukon ne sont pas disponibles avant 2023 ».

De plus, les données couplées ne peuvent pas être utilisées pour produire des statistiques liées aux établissements déclarés dans la Base de données sur les congés des patients (BDGP), le Système national d'information sur les soins ambulatoires (SNISA) et le Système d'information ontarien sur la santé mentale (SIOSM), et tout produit de données à l'échelle de l'établissement doit faire l'objet d'une utilisation restreinte conformément aux lignes directrices relatives au contrôle de la confidentialité. Les renseignements sur l'établissement institutionnel, comme le code postal figurant dans la BDGP, le SNISA et le SIOSM, peuvent être utilisés pour obtenir des renseignements (p. ex. distance jusqu'à un établissement de soins de courte durée), mais ils ne peuvent pas être utilisés comme résultat d'intérêt.

Annexe A – Répartition en pourcentage de certaines caractéristiques démographiques, sociales et économiques de la Cohorte santé et environnement du recensement canadien de 2021 et du fichier du Centre de données de recherche du Recensement de 2021

Le Tableau 10 présente la répartition en pourcentage de certaines caractéristiques démographiques, sociales et économiques de la Cohorte santé et environnement du recensement canadien de 2021 et du fichier du Centre de données de recherche du Recensement de 2021.

Tableau 10

Répartition en pourcentage de certaines caractéristiques démographiques, sociales et économiques de la Cohorte santé et environnement du recensement canadien (CSERCCan) de 2021 et du fichier du Centre de données de recherche du Recensement de 2021

Caractéristiques de la population	CSERCCan de 2021		Recensement de 2021
	Non pondérée	Pondérée	Pondérée
	nombre (arrondi)		
	8 660 375	36 328 475	36 328 475
	pourcentage de la taille		
Taille			
Genre¹			
Femme+ ²	50,7	50,6	50,6
Homme+ ³	49,3	49,4	49,4
Âge par groupes			
0 à 14 ans	16,8	16,5	16,5
15 à 24 ans	11,4	11,5	11,5
25 à 34 ans	13,3	13,5	13,5
35 à 44 ans	13,4	13,4	13,4
45 à 54 ans	12,8	12,8	12,8
55 à 64 ans	14,2	14,2	14,2
65 à 74 ans	11,0	11,0	11,0
75 à 84 ans	5,4	5,4	5,4
85 ans et plus	1,7	1,7	1,7
État matrimonial (de facto)			
Marié ou vivant en union libre	48,6	48,4	48,2
Ne vivant pas en union libre (jamais marié, séparé, divorcé ou veuf)	51,4	51,6	51,8
Mobilité : Lieu de résidence 1 an auparavant (2020)			
Même adresse (logement)	87,5	87,0	87,0
Adresse différente (logement)	10,8	11,3	11,3
À l'extérieur du Canada	0,7	0,8	0,8
Sans objet (exclus en raison de l'âge)	1,0	0,9	0,9
Scolarité : Plus haut certificat, diplôme ou grade			
Aucun certificat, diplôme ou grade	13,7	13,4	13,5
Diplôme d'études secondaires ou attestation d'équivalence	22,0	22,2	22,3
Certificat ou diplôme d'apprenti ou d'une école de métiers	7,3	7,3	7,3
Certificat ou diplôme d'un collège, d'un cégep ou d'un autre établissement non universitaire	15,7	15,8	15,7
Certificat ou diplôme universitaire inférieur au baccalauréat	2,4	2,5	2,5
Baccalauréat ou grade supérieur	22,1	22,4	22,3
Sans objet (< 15 ans)	16,8	16,5	16,5
Situation d'activité pendant la semaine du dimanche 2 mai au samedi 8 mai 2021			
Personnes occupées	47,5	47,7	47,7
Chômeurs	5,4	5,5	5,5
Inactifs	30,3	30,3	30,3
Sans objet (< 15 ans)	16,8	16,5	16,5
Situation de faible revenu fondé sur la mesure de faible revenu après impôt			
Pas à faible revenu	88,9	89,0	88,9
À faible revenu	11,1	11,0	11,1

Tableau 10
Répartition en pourcentage de certaines caractéristiques démographiques, sociales et économiques de la Cohorte santé et environnement du recensement canadien (CSERCan) de 2021 et du fichier du Centre de données de recherche du Recensement de 2021

Caractéristiques de la population	CSERCan de 2021		Recensement de 2021
	Non pondérée	Pondérée	Pondérée
Identité autochtone			
Premières Nations (Indiens de l'Amérique du Nord)	4,4	2,9	2,9
Métis	1,7	1,7	1,7
Inuk (Inuit)	0,5	0,2	0,2
Réponses autochtones multiples	0,1	0,1	0,1
Réponses autochtones non comprises ailleurs	0,1	0,1	0,1
Identité non autochtone	93,3	95,0	95,0

1. Étant donné que la taille de la population non binaire est petite, il est parfois nécessaire d'agréger les données dans une variable sur le genre à deux catégories pour protéger la confidentialité des réponses. Dans ces cas, les personnes dans la catégorie « personnes non binaires » sont réparties dans les deux autres catégories de genre et sont désignées par le signe +.

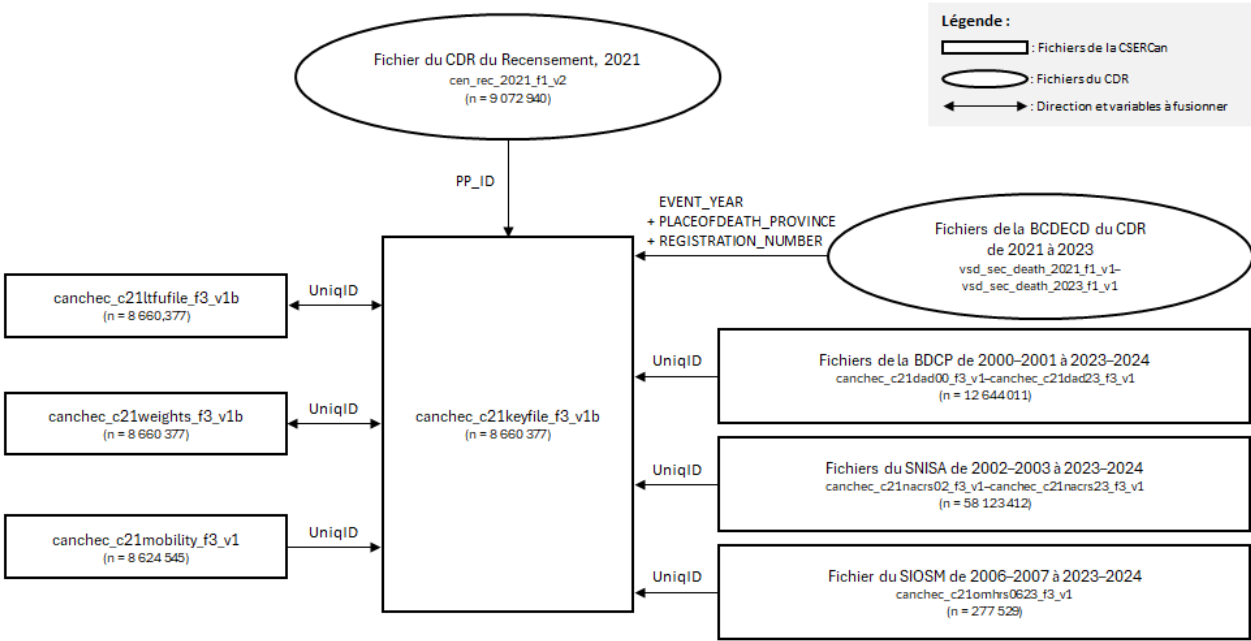
2. Cette catégorie comprend les hommes (et/ou les garçons) de même que certaines personnes non binaires.

3. Cette catégorie comprend les femmes (et/ou les filles) de même que certaines personnes non binaires.

Sources : Statistique Canada, Cohorte santé et environnement du recensement canadien de 2021 (numéro d'enregistrement 5422), Recensement de la population, 2021 (numéro d'enregistrement 3901) et Base canadienne de données de l'état civil – Décès, de 2021 à 2023 (numéro d'enregistrement 3233).

Annexe B – Schéma de la base de données de la Cohorte santé et environnement du recensement canadien de 2021 pour les utilisateurs des données des Centres de données de recherche

Figure 1
Schéma de la base de données de la Cohorte santé et environnement du recensement canadien de 2021 pour les utilisateurs des données des Centres de données de recherche



Notes : Que les objets ne sont pas dessinés à l'échelle.

Acronymes et abréviations: CSERCan: Cohortes santé et environnement du recensement canadien; BCDECD: Base canadienne de données de l'état civil — Décès; REC: Recensement de la population; BDCP: Base de données sur les congés des patients; SNISA: Système national d'information sur les soins ambulatoires; SIOSM: Système d'information ontarien sur la santé mentale; CDR: Centres de données de recherche.

La version 1b de la CSERCan de 2021 a été préparée en utilisant cen_rec_2021_f1_v2 et vsd_sec_death_2021_f1_v1-vsd_sec_death_2023_f1_v1; toutefois, d'autres versions du Recensement de la population de 2021 et de la BCDECD peuvent être utilisées avec les clés de couplage de la version 1b de la CSERCan de 2021.

Sources : Statistique Canada, Cohorte santé et environnement du recensement canadien de 2021 (numéro d'enregistrement 5422), Recensement de la population, 2021 (numéro d'enregistrement 3901) et Base canadienne de données de l'état civil – Décès, de 2021 à 2023 (numéro d'enregistrement 3233); Agence du revenu du Canada, Fichier maître des particuliers T1; et Institut canadien d'information sur la santé, Base de données sur les congés des patients, de 2000–2001 à 2023–2024(métadonnées), Système national d'information sur les soins ambulatoires, de 2002–2003 à 2023–2024 (métadonnées) et Système d'information ontarien sur la santé mentale, de 2006–2007 à 2023–2024 (métadonnées).