

Mesurer les inégalités en santé : trousse d'outils

Calcul des taux stratifiés et des mesures
de l'inégalité : méthodologie et codes
en SAS et en R



Institut canadien
d'information sur la santé

Canadian Institute
for Health Information

La production du présent document est rendue possible grâce à un apport financier de Santé Canada et des gouvernements provinciaux et territoriaux. Les opinions exprimées dans ce rapport ne représentent pas nécessairement celles de Santé Canada ou celles des gouvernements provinciaux et territoriaux.

Tous droits réservés.

Le contenu de cette publication peut être reproduit tel quel, en tout ou en partie et par quelque moyen que ce soit, uniquement à des fins non commerciales pourvu que l'Institut canadien d'information sur la santé soit clairement identifié comme le titulaire du droit d'auteur. Toute reproduction ou utilisation de cette publication et de son contenu à des fins commerciales requiert l'autorisation écrite préalable de l'Institut canadien d'information sur la santé. La reproduction ou l'utilisation de cette publication ou de son contenu qui sous-entend le consentement de l'Institut canadien d'information sur la santé, ou toute affiliation avec celui-ci, est interdite.

Pour obtenir une autorisation ou des renseignements, veuillez contacter l'ICIS :

Institut canadien d'information sur la santé

495, chemin Richmond, bureau 600

Ottawa (Ontario) K2A 4H6

Téléphone : 613-241-7860

Télécopieur : 613-241-8120

icis.ca

droitdauteur@icis.ca

ISBN 978-1-77479-111-0 (PDF)

© 2022 Institut canadien d'information sur la santé

Comment citer ce document :

Institut canadien d'information sur la santé. *Mesurer les inégalités en santé : trousse d'outils — calcul des taux stratifiés et des mesures de l'inégalité : méthodologie et codes en SAS et en R*. Ottawa, ON : ICIS; 2022.

This publication is also available in English under the title *Measuring Health Inequalities: A Toolkit — Calculating Stratified Rates and Inequality Measures: Methodology and Code in SAS and R*.

ISBN 978-1-77479-110-3 (PDF)

Table des matières

Conditions d'utilisation des codes en SAS et en R.....	4
Introduction.....	6
Notes méthodologiques.....	6
Taux normalisés selon l'âge.....	6
Mesures simples de l'inégalité.....	7
Mesures complexes de l'inégalité.....	8
Calcul et interprétation de la RPT et du NPRP.....	11
Points à considérer lors de la déclaration des mesures sommaires.....	14
Annexes.....	15
Annexe A : aperçu des codes.....	15
Annexe B : Macro SAS — %Calculate_Stratified_Rates.....	17
Annexe C : Macro SAS — %Calculate_Inequality_Measures.....	29
Annexe D : Fonction R — Calculate_Stratified_Rates.....	39
Annexe E : Fonction R — Calculate_Inequality_Measures.....	48
Références.....	53

Conditions d'utilisation des codes en SAS et en R

Vous pouvez reproduire, modifier, distribuer et utiliser sans frais les codes en SAS et en R qui figurent aux annexes B, C, D et E de la présente publication (les **Codes**) à des fins non commerciales, notamment dans le cadre d'études et de recherches universitaires et gouvernementales, pour le renforcement des capacités de production de rapports sur l'équité au Canada, à condition

- d'identifier clairement l'ICIS comme le titulaire du droit d'auteur des Codes;
- d'inclure ces conditions d'utilisation dans toute reproduction intégrale ou partielle des Codes, y compris toute version modifiée des Codes;
- d'indiquer clairement toute modification que vous avez apportée aux Codes;
- de ne pas utiliser les Codes à des fins commerciales (p. ex. les intégrer à un produit en vue de sa vente);
- de ne pas énoncer ou laisser entendre que l'ICIS utilise, approuve ou soutient tout produit ou service ou toute cause, opinion, personne ou entité, ou qu'il y est affilié.

Vous reconnaissez et convenez par ailleurs de ce qui suit :

- Le non-respect de ces conditions d'utilisation peut entraîner des conséquences comprenant, entre autres, la révocation des droits qui vous sont accordés en vertu de ces conditions d'utilisation et l'obligation de détruire ou de restituer les Codes et toute reproduction intégrale ou partielle de ceux-ci.
- Vous ferez preuve de prudence ou vous vous abstenrez même de publier des données générées à l'aide des Codes si ces données identifient ou pourraient identifier une personne ou un établissement de soins de santé.
- Ces conditions d'utilisation constituent l'entente intégrale en ce qui a trait aux Codes.
- L'ICIS détient les droits et titres relatifs aux Codes. Aucun droit de propriété ni aucun autre droit ne vous est accordé, sauf dans les cas mentionnés ci-dessus.
- Vous ne prendrez aucune mesure susceptible de compromettre les droits de propriété intellectuelle de l'ICIS sur les Codes, ou l'application de ces conditions d'utilisation aux Codes.

- L'ICIS FOURNIT LES CODES « TELS QUELS » SANS GARANTIE DE QUELQUE NATURE QUE CE SOIT, EXPRESSÉMENT OU IMPLICITEMENT, CONFORMÉMENT À LA LOI OU DE TOUTE AUTRE MANIÈRE, NOTAMMENT RELATIVEMENT À LA QUALITÉ MARCHANDE, À L'ADAPTATION À UN USAGE PARTICULIER ET À L'ABSENCE DE CONTREFAÇON. L'ICIS NE SERA EN AUCUN CAS TENU RESPONSABLE DE QUELQUE PERTES OU DOMMAGES, QUELS QU'ILS SOIENT, QUI POURRAIENT DÉCOULER DES CODES OU S'Y RAPPORTER.
- VOUS DÉGAGEREZ L'ICIS DE TOUTE RESPONSABILITÉ EN CAS DE PERTES OU DE DOMMAGES DÉCOULANT DE VOTRE UTILISATION DES CODES.
- Ces conditions d'utilisation sont régies par les lois applicables dans la province de l'Ontario et au Canada, et doivent être interprétées en conséquence, sans égard aux principes de conflits de lois.

Vous confirmez que vous comprenez et acceptez ces conditions d'utilisation lors de votre première utilisation des Codes. SI VOUS N'ÊTES PAS EN MESURE D'ACCEPTER CES CONDITIONS D'UTILISATION, VOUS NE POUVEZ PAS UTILISER LES CODES.

Introduction

Le présent guide donne un aperçu des codes en SAS et en R produits par l'Institut canadien d'information sur la santé (ICIS) pour le calcul des taux des indicateurs de santé bruts et normalisés selon l'âge, stratifiés selon différents facteurs (quintile de revenu, emplacement géographique [région urbaine ou rurale/éloignée], etc.), et pour le calcul de 4 mesures sommaires de l'inégalité : ratio des taux, différence des taux, réduction potentielle du taux et nombre de personnes à risque au sein de la population. La méthodologie décrit l'approche utilisée pour le calcul des taux et des inégalités et s'applique aux codes en SAS et en R.

Notes méthodologiques

Taux normalisés selon l'âge

L'ICIS calcule les taux des indicateurs normalisés selon l'âge au moyen d'une méthode de normalisation directe qui utilise la population canadienne de 2011 (selon le Recensement de 2011) comme population de référence. La normalisation est effectuée en fonction de groupes d'âge de 5 ans.

Le **taux normalisé selon l'âge** correspond à la somme de tous les taux pondérés selon l'âge :

$$\text{Taux normalisé selon l'âge} = \sum \text{Taux pondérés selon l'âge}$$

Les **taux pondérés selon l'âge** sont calculés pour chaque groupe d'âge au moyen de la formule suivante :

$$\text{Taux pondéré selon l'âge} = \text{Taux}_{\text{brut}} \times \text{Pondération} \times \text{Multiplicateur}$$

où le $\text{Taux}_{\text{brut}}$ pour chaque groupe d'âge correspond au quotient du nombre de cas lié à l'indicateur divisé par la population, la **Pondération** est la valeur de pondération de chaque groupe d'âge de la population de référence, et le **Multiplicateur** est la valeur pour laquelle vous voulez présenter les taux (p. ex. par 100 000 habitants).

La **variance** du taux pondéré selon l'âge est calculée pour chaque groupe d'âge selon une distribution binomiale au moyen de la formule suivante :

$$\text{Variance (Taux pondéré selon l'âge)} = \text{Pondération}^2 \times \text{Taux}_{\text{brut}} \times \frac{(\text{Multiplicateur} - (\text{Taux}_{\text{brut}}))}{\text{Population dans le groupe d'âge}}$$

Ensuite, la variance du taux normalisé selon l'âge correspond à la somme des variances pour tous les taux pondérés selon l'âge :

$$\text{Variance (Taux normalisé selon l'âge)} = \sum \text{Variance (Taux pondérés selon l'âge)}$$

Mesures simples de l'inégalité

Le **ratio des taux (RT)** est une mesure de l'inégalité relative entre sous-groupes. On le calcule en divisant le taux du groupe de comparaison par le taux du groupe de référence.

$$\text{Ratio des taux}_{\text{comparaison}} = \text{Taux}_{\text{comparaison}} \div \text{Taux}_{\text{référence}}$$

Exemple : Mesurer les inégalités liées au revenu, où le premier quintile (Q1) est le quintile de revenu le plus faible (c.-à-d. le groupe de comparaison) et le cinquième quintile (Q5), le quintile de revenu le plus élevé (c.-à-d. le groupe de référence). Le taux dans le Q1 est de 10 cas par 100 000 et le taux dans le Q5, de 5 cas par 100 000.

$$\text{Taux}_{\text{Q1}} \div \text{Taux}_{\text{Q5}} = 10 \text{ par } 100\,000 \div 5 \text{ par } 100\,000 = 2$$

Interprétation : Le taux de l'affection X est 2 fois plus élevé chez les Canadiens du quintile de revenu le plus faible que chez ceux du quintile de revenu le plus élevé.

La **variance** du logarithme du ($\text{RT}_{\text{comparaison}}$) est calculée au moyen de la formule suivante^{1, i} :

$$\text{Variance}(\text{Logarithme}(\text{RT}_{\text{comparaison}})) = \frac{\text{Variance}(\text{Taux}_{\text{comparaison}})}{\text{Taux}_{\text{comparaison}}^2} + \frac{\text{Variance}(\text{Taux}_{\text{référence}})}{\text{Taux}_{\text{référence}}^2}$$

L'**intervalle de confiance de 95 % du $\text{RT}_{\text{comparaison}}$** est calculé au moyen de la formule

$$e^{(\text{Logarithme}(\text{RT}_{\text{comparaison}}) \pm 1,96 \sqrt{\text{Variance}(\text{Logarithme}(\text{RT}_{\text{comparaison}})))}$$

La **différence des taux (DT)** est une mesure de l'inégalité *absolue* entre sous-groupes. On la calcule en soustrayant le taux du groupe de référence du taux du groupe de comparaison.

Exemple : $\text{Taux}_{\text{Q1}} - \text{Taux}_{\text{Q5}} = 10 \text{ par } 100\,000 - 5 \text{ par } 100\,000 = 5 \text{ par } 100\,000$

Interprétation : L'affection X atteint 5 Canadiens de plus par 100 000 habitants dans le quintile de revenu le plus faible que dans le quintile de revenu le plus élevé.

i. Les formules servant au calcul des variances et des intervalles de confiance de 95 % sont fournies pour les données administratives et pourraient devoir être modifiées selon la source des données (p. ex. pour les données d'enquête, il faudra tenir compte de la stratégie d'échantillonnage).

La **variance** de la différence des taux est calculée au moyen de la formule suivante :

$$\text{Variance (DT}_{\text{comparaison}}) = \text{Variance (Taux}_{\text{comparaison}}) + \text{Variance (Taux}_{\text{référence}})$$

L'**intervalle de confiance de 95 % de la DT**_{comparaison} est calculé au moyen de la formule **DT**_{comparaison} \pm 1,96 $\sqrt{\text{Variance (DT}_{\text{comparaison}})}$.

Mesures complexes de l'inégalité

La réduction potentielle du taux et le nombre de personnes à risque au sein de la population sont 2 exemples de mesures complexes de l'inégalité. Ce type de mesure peut tenir compte des données de tous les sous-groupes de populationⁱⁱ (p. ex. inégalité à l'échelle de tous les quintiles de revenu) et donne comme résultat une valeur unique qui rend compte du niveau d'inégalité.

La **réduction potentielle du taux (RPT)** est une mesure *relative* de la réduction potentielle du taux d'un indicateur de santé si chaque sous-groupe de population affichait le même taux que le sous-groupe ayant le taux le plus souhaitable. Elle est aussi communément appelée **fraction attribuable dans la population** ou **risque attribuable à la population**². Vous pouvez utiliser la RPT lorsque de faibles taux sont recherchés pour l'indicateur et que le groupe de référence est le sous-groupe affichant le taux le plus souhaitable.

Exemple : RPT = 45 %

Interprétation : Au cours d'une année donnée, 45 % des hospitalisations liées à l'affection X auraient pu être évitées si les Canadiens de tous les quintiles de revenu avaient affiché le même taux d'hospitalisations que ceux du quintile de revenu le plus élevé. (Voir [Calcul et interprétation de la RPT et du NPRP.](#))

La RPT et son **intervalle de confiance de 95 %** sont calculés au moyen de la formule ci-dessous pour un facteur de stratification comportant un nombre « n » de catégories (p. ex. pour un facteur de stratification à 5 catégories tel que le quintile de revenu, n = 5). La RPT est habituellement présentée sous forme de pourcentage, comme ceci :

$$\text{RPT} = \frac{\sum_{i=1}^n P_i \left(\frac{\text{Taux}_i}{\text{Taux}_{\text{référence}}} - 1 \right)}{1 + \sum_{i=1}^n P_i \left(\frac{\text{Taux}_i}{\text{Taux}_{\text{référence}}} - 1 \right)} \times 100 \%$$

où P_i est la proportion de la population dans la catégorie i .

ii. Il existe d'autres façons de calculer la réduction potentielle du taux (aussi appelée fraction attribuable dans la population) pour mesurer les inégalités en santé. Par exemple, dans l'Outil de données sur les inégalités en santé de l'Agence de la santé publique du Canada, les fractions attribuables dans la population sont calculées pour chaque groupe de comparaison, au lieu de servir de valeur unique qui rend compte de l'inégalité à l'échelle de tous les sous-groupes de population.

La limite de confiance inférieure (LCI) et la limite de confiance supérieure (LCS) de la RPT sont ensuite obtenues comme suit :

$$LCI_{RPT} = 1 - \frac{1}{P_n + e^{(\text{Logarithme (Ratio des taux}_{1:(n-1), n}) - 1,96\sqrt{\text{Variance (Ratio des taux}_{1:(n-1), n})})}}$$

$$LCS_{RPT} = 1 - \frac{1}{P_n + e^{(\text{Logarithme (Ratio des taux}_{1:(n-1), n}) + 1,96\sqrt{\text{Variance (Ratio des taux}_{1:(n-1), n})})}}$$

où Ratio des taux_{1:(n-1), n}

est le ratio entre la somme de la proportion de la population (P_i), multipliée par le taux normalisé selon l'âge (Taux_{*i*}) de la première catégorie ($n - 1$) et le taux de la catégorie de référence, calculé comme suit :

$$\frac{\sum_{i=1}^{(n-1)} P_i \text{Taux}_i}{\text{Taux}_{\text{référence}}}$$

avec la variance

$$\frac{\sum_{i=1}^{(n-1)} P_i^2 \times \text{Variance (Taux}_i)}{(\sum_{i=1}^{(n-1)} P_i \text{Taux}_i)^2} + \frac{\text{Variance (Taux}_{\text{référence}})}{\text{Taux}_{\text{référence}}^2}$$

Le nombre de personnes à risque au sein de la population (NPRP) est une mesure *absolue* de la réduction potentielle du nombre de cas associés à un indicateur de santé si chaque sous-groupe de population affichait le même taux que le sous-groupe ayant le taux le plus souhaitable. Cette mesure rend compte du gradient de l'inégalité à l'échelle de plusieurs catégories, comme les quintiles de revenu.

Exemple : NPRP = 18 700

Interprétation : Au cours d'une année donnée, environ 18 700 hospitalisations liées à l'affection X auraient pu être évitées si les Canadiens de tous les niveaux de revenu avaient affiché le même taux d'hospitalisations que ceux du niveau de revenu le plus élevé. (Voir [Calcul et interprétation de la RPT et du NPRP.](#))

Le NPRP est lié à la RPT de la manière suivante :

Nombre de personnes à risque au sein de la population = Taux global normalisé
de l'indicateur $\times N_{\text{population totale}} \times \text{RPT}$

L'**amélioration potentielle du taux (APT)** est une mesure de l'inégalité *relative* (analogue à la RPT) utilisée lorsque des taux élevés sont souhaitables pour l'indicateur. Elle est aussi communément appelée fraction préventive.

L'APT est liée à la RPT de la manière suivante :

$$APT = 1 - \frac{1}{1 - RPT}$$

où la RPT désigne la réduction potentielle du taux exprimée sous forme de décimale.

La variance de l'APT est déterminée d'après le calcul de la variance de la RPT au moyen de l'équation suivante :

$$Variance(APT) = Variance\left(1 - \frac{1}{1 - RPT}\right)$$

$$Variance(APT) = Variance\left(\text{Logarithme}\left(\frac{Taux_{1:(n-1)}}{Taux_{référence}}\right)\right) = \frac{Variance(Taux_{référence})}{(Taux_{référence})^2} + \frac{\sum_{i=1}^{(n-1)} P_i^2 Variance(Taux_i)}{(Taux_{1:(n-1)})^2}$$

Les limites inférieure et supérieure de l'intervalle de confiance de 95 % (LCI et LCS) sont donc calculées comme suit :

$$LCI_{APT} = \text{Exp}(\text{Logarithme}(APT) - 1,96\sqrt{Variance}(APT))$$

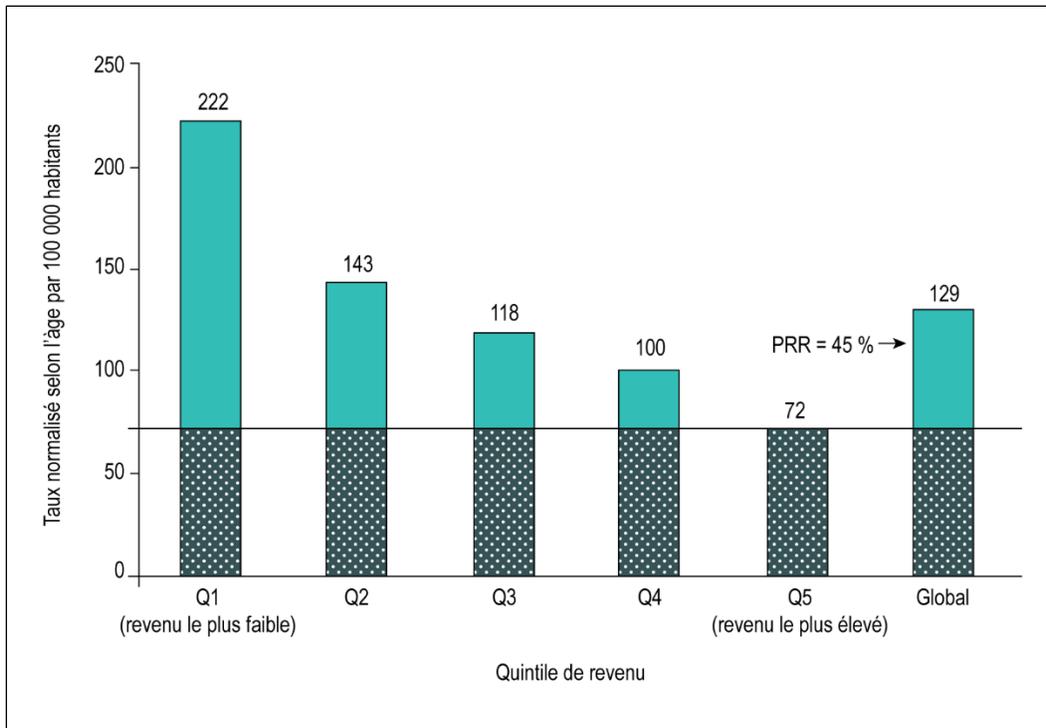
$$LCS_{APT} = \text{Exp}(\text{Logarithme}(APT) + 1,96\sqrt{Variance}(APT))$$

Calcul et interprétation de la RPT et du NPRP

Examinons un exemple à l'aide de l'indicateur Hospitalisations liées à la maladie pulmonaire obstructive chronique (MPOC) chez les Canadiens de moins de 75 ans afin de comprendre comment calculer et interpréter la RPT et le NPRP.

Les taux d'hospitalisations normalisés selon l'âge de cet indicateur, lorsqu'ils sont stratifiés par quintile de revenu, varient de 72 par 100 000 habitants au niveau de revenu le plus élevé (Q5) à 222 par 100 000 habitants au niveau de revenu le plus faible (Q1).

Figure Taux d'hospitalisations liées à la MPOC chez les Canadiens de moins de 75 ans, selon le revenu, 2012



Calcul de la RPT

Lorsque nous calculons la RPT, nous cherchons à estimer de combien inférieur serait le taux global si tous les quintiles de revenu affichaient le même taux que le groupe ayant le taux le plus souhaitable (c.-à-d. le quintile de revenu le plus élevé).

Dans la figure, la ligne horizontale représente le taux de l'indicateur au niveau de revenu le plus élevé (Q5).

Pour calculer la RPT, vous avez besoin des éléments d'information suivants :

- le taux normalisé selon l'âge pour chaque quintile de revenu ($Taux_i$)
- la proportion de la population dans chaque quintile de revenu (P_i) (voir l'encadré ci-dessous)

Insérez les valeurs dans la formule suivante de la RPT :

$$RPT = \frac{\sum_{i=1}^5 P_i \left(\frac{Taux_i}{Taux_5} - 1 \right)}{1 + \sum_{i=1}^5 P_i \left(\frac{Taux_i}{Taux_5} - 1 \right)} \times 100 \%$$

$$RPT = \frac{\left[0,20 \left(\frac{222}{72} - 1 \right) + 0,20 \left(\frac{143}{72} - 1 \right) + 0,19 \left(\frac{118}{72} - 1 \right) + 0,21 \left(\frac{100}{72} - 1 \right) + 0,20 \left(\frac{72}{72} - 1 \right) \right]}{1 + \left[0,20 \left(\frac{222}{72} - 1 \right) + 0,20 \left(\frac{143}{72} - 1 \right) + 0,19 \left(\frac{118}{72} - 1 \right) + 0,21 \left(\frac{100}{72} - 1 \right) + 0,20 \left(\frac{72}{72} - 1 \right) \right]} \times 100 \%$$

$$RPT = 45 \%$$

Interprétation de la RPT

En 2012, environ 45 % des hospitalisations liées à la MPOC auraient pu être évitées chez les Canadiens de moins de 75 ans si tous les quintiles de revenu avaient affiché le même faible taux d'hospitalisations liées à la MPOC que le quintile de revenu le plus élevé.

Encadré : Comprendre la proportion de la population dans chaque quintile de revenu (P_i)

Par définition, la proportion de la population dans chaque quintile est de 20 %. Cependant, dans le cas de l'indicateur sur les hospitalisations liées à la MPOC, la proportion de la population dans chaque quintile de revenu n'est pas exactement de 20 %.

Pour cet indicateur, la proportion de la population dans chaque quintile de revenu est en effet tirée des estimations démographiques censitaires pour les Canadiens de moins de 75 ans (c.-à-d. le dénominateur pour chaque quintile divisé par la population totale âgée de 0 à 74 ans).

Ces estimations ont été réparties dans des quintiles de revenu établis en fonction de toute la population canadienne (tous âges confondus). Par conséquent, elles reflètent le revenu de tous les Canadiens, et non le revenu des Canadiens de moins de 75 ans seulement.

Dans le calcul ci-dessus de la RPT, la proportion de la population dans chaque quintile de revenu est la suivante : $P_1 = 0,20$, $P_2 = 0,20$, $P_3 = 0,19$, $P_4 = 0,21$ et $P_5 = 0,20$.

Calcul du NPRP

Lorsque nous calculons le NPRP, nous convertissons la RPT en nombre approximatif de cas qui pourraient être évités si le taux de tous les sous-groupes de population était le même que celui du sous-groupe ayant le taux le plus souhaitable.

Pour calculer le NPRP, vous avez besoin des éléments d'information suivants :

- le taux global de l'indicateur
- la RPT
- la population d'intérêt totale (dans cette analyse, la population de Canadiens âgés de 0 à 74 ans en 2012 est de 32 079 232)

Insérez les valeurs dans la formule suivante du NPRP :

$$\begin{aligned} \text{NPRP} &= \text{Taux global normalisé de l'indicateur} \times N_{\text{population totale}} \times \text{RPT} \\ &= 129 \div 100\,000 \times 32\,079\,232 \times 45 \div 100 \end{aligned}$$

$$\text{NPRP} = 18\,663$$

Interprétation du NPRP

En 2012, il y aurait eu environ 18 700 hospitalisations de moins liées à la MPOC chez les Canadiens de moins de 75 ans si tous les quintiles de revenu avaient affiché un taux aussi faible que celui du quintile de revenu le plus élevé.

Remarque : Bien que la valeur calculée soit de 18 663, nous déclarons le NPRP comme un nombre approximatif (arrondi à la centaine près dans le présent exemple).

Points à considérer lors de la déclaration des mesures sommaires

L'ICIS utilise un intervalle de confiance de 95 % pour déterminer la signification statistique d'une mesure sommaire de l'inégalité en santé, comme l'illustre le tableau suivant (à noter qu'il est possible d'utiliser une autre approche).

Mesure de l'inégalité	Considérée comme statistiquement significative si l'intervalle de confiance de 95 %
Ratio des taux (RT)	Ne comprend pas 1
Différence des taux (DT)	Ne comprend pas 0
Réduction potentielle du taux (RPT)	Ne comprend pas 0
Nombre de personnes à risque au sein de la population (NPRP)	Ne comprend pas 0
Amélioration potentielle du taux (APT)	Ne comprend pas 0

- Voici les pratiques exemplaires de l'ICIS pour la déclaration des mesures sommaires de l'inégalité :
 - Déclarer la RPT et le NPRP comme nuls si la RPT est négative et statistiquement différente de 0 (c.-à-d. que les limites de confiance inférieure et supérieure sont négatives).
 - Supprimer et déclarer comme nulles les valeurs d'APT qui sont négatives et statistiquement significatives.
 - Déclarer le NPRP comme un nombre approximatif (p. ex. arrondi à la centaine près).

Annexes

Annexe A : aperçu des codes

Calculate_Stratified_Rates

Ce code calcule les taux bruts et normalisés selon l'âge, stratifiés selon le facteur de stratification de l'équité de votre choix (quintile de revenu, région urbaine ou rurale/éloignée, etc.). Les taux stratifiés sont aussi calculés en fonction du sexe (hommes, femmes ou les 2 sexes combinés) et du niveau de déclaration (à l'échelle du Canada ou des provinces et territoires). Ce code peut être modifié pour le calcul des taux stratifiés de différentes autorités compétentes.

Pour utiliser le code

- Créez les 3 fichiers d'entrée suivants, formatés selon la convention décrite dans le fichier Excel [Mesurer les inégalités en santé : trousse d'outils — formats des fichiers d'entrée pour les macros SAS et les fonctions R](#).
 1. Cas liés à l'indicateur
 2. Estimations de la population
 3. Estimations de la population de référence

Le fichier Excel contient des fichiers de données fictives pour les 3 types de fichiers d'entrée. Vous pouvez les utiliser pour exécuter des exemples d'analyses.

- Définissez tous les paramètres d'entrée (p. ex. tranche d'âge pour l'indicateur).
- Lancez le code Calculate_Stratified_Rates (consultez l'[annexe B](#) pour la macro SAS et l'[annexe D](#) pour la fonction R).
- Vérifiez que les données générées sont complètes, sans erreur ni avertissement.

Il faut utiliser les types de variables (type numérique ou type caractère) indiqués dans le fichier Excel. Les autres exigences relatives au format présentées dans le fichier Excel s'appliquent uniquement aux macros SAS.

Calculate_Inequality_Measures

Ce code calcule les mesures de l'inégalité, dont le ratio des taux, la différence des taux, la réduction potentielle du taux et le nombre de personnes à risque au sein de la population, pour des indicateurs de santé stratifiés par facteur de stratification de l'équité (p. ex. le quintile de revenu). Les résultats sont générés en fonction du sexe (hommes, femmes ou les 2 sexes combinés) et du niveau de déclaration (à l'échelle du Canada ou des provinces et territoires).

Pour utiliser le code

- Créez un fichier d'entrée formaté selon la convention décrite dans le fichier Excel [Mesurer les inégalités en santé : trousse d'outils — formats des fichiers d'entrée pour les macros SAS et les fonctions R](#). Notez que les données générées après l'exécution réussie du code Calculate_Stratified_Rates seront correctement formatées.
- Lancez le code Calculate_Inequality_Measures (consultez l'[annexe C](#) pour la macro SAS et l'[annexe E](#) pour la fonction R).
- Vérifiez que les données générées sont complètes, sans erreur ni avertissement.

Annexe B : Macro SAS — %Calculate_Stratified_Rates

Fichier de sortie

&indicator._&YR._&AGEGPL._&AGEGPU._&equity_stratifier._rates — Taux bruts et normalisés selon l'âge pour le facteur de stratification de l'équité et le niveau de déclaration choisis.

Paramètres

Vous devez définir tous les paramètres ci-dessous.

- *indicator* — Nom de l'indicateur de santé mesuré. Il peut s'agir de n'importe quel nom de votre choix.
- *yr* — Année. Les données sur les cas liés à l'indicateur et les estimations de la population seront uniquement extraites pour l'année indiquée. Une seule année peut être saisie à la fois.
- *reporting_level* — Niveau de déclaration géographique. La macro calcule les taux à l'échelle du Canada et des provinces et territoires. Donnez à la variable *reporting_level* le même nom dans votre fichier des cas liés à l'indicateur (*infile*) et votre fichier des estimations de la population (*popfile*).
- *equity_stratifier* — Variable pour laquelle vous désirez calculer les taux stratifiés (p. ex. le quintile de revenu). Il peut s'agir d'un facteur de stratification ordonné, tel que le quintile de revenu, ou d'un facteur de stratification non ordonné, tel que l'emplacement géographique (région urbaine ou région rurale/éloignée). Vous devez utiliser le même nom pour la variable *equity_stratifier* dans tous vos fichiers d'entrée.
- *inlib* — Nom du répertoire SAS où tous vos fichiers d'entrée sont stockés.
- *infile* — Nom du fichier d'entrée des cas liés à l'indicateur.
- *stdpop* — Nom du fichier des estimations de la population de référence.
- *popfile* — Nom du fichier des estimations de la population catégorisées selon le facteur de stratification de l'équité choisi (p. ex. estimations de la population par quintile de revenu).
- *popvar* — Chiffres concernant la population provenant du fichier des estimations de la population (*popfile*).
- *age_var* — Nom de la variable de l'âge utilisée uniformément dans le fichier des cas liés à l'indicateur (*infile*), le fichier des estimations de la population (*popfile*) et le fichier des estimations de la population de référence (*stdpop*).
- *agegpl* — Limite d'âge inférieure de l'indicateur de santé; doit respecter le format 01, 02, [...], 14, 15 tel que précisé dans la variable de l'âge (p. ex. si la limite d'âge inférieure pour votre indicateur de santé est 0 an, cette valeur doit être 01).

- *agegpu* — Limite d'âge supérieure de l'indicateur de santé; doit respecter le format 01, 02, [...], 14, 15 tel que précisé dans la variable de l'âge (p. ex. si la limite d'âge supérieure pour votre indicateur de santé est 74 ans, cette valeur doit être 15).
- *multiplier* — Valeur pour laquelle vous voulez présenter les taux (p. ex. MULTIPLIER = 100,000 pour les taux par 100 000).
- *roundunit* — Valeur d'arrondissement (p. ex. pour arrondir à une décimale, précisez ROUNDUNIT = .01).
- *outlib* — Nom du répertoire SAS où vous voulez stocker votre fichier de sortie.
- *outfile* — Nom du fichier de sortie.

Invocation de la macro

%MACRO Calculate_Stratified_Rates (indicator=, yr=, reporting_level=, equity_stratifier=, inlib=, infile=, stdpop=, popfile=, popvar=, age_var=, agegpu=, multiplier=, roundunit=, outlib=, outfile=);

L'exemple ci-dessous montre comment la macro peut être invoquée si l'on utilise les données fictives fournies dans le fichier Excel. Dans cet exemple, on suppose que tous les fichiers sont conservés dans le répertoire *work*.

```
%Calculate_Stratified_Rates (  
    indicator=fake,  
    yr=2020,  
    reporting_level=PT_code,  
    equity_stratifier=income_quintile,  
    inlib=work,  
    infile=Example_infile_data,  
    stdpop=work.Example_stdpop_file,  
    popfile=work.Example_popfile_data,  
    popvar=PT_pop,  
    age_var=AGE_GROUP_CODE,  
    agegpu=19,  
    multiplier=100000,  
    roundunit=.01,  
    outlib=work,  
    outfile=fakedata_output1);
```

Étapes de la macro

1. Prépare les cas liés à l'indicateur en créant des données agrégées par niveau de déclaration, groupe d'âge, sexe et facteur de stratification de l'équité.
2. Prépare le fichier de population de référence qui servira à la normalisation selon l'âge.
3. Prépare les dénominateurs de population de la tranche d'âge visée.
4. Lie les 3 fichiers de données ci-dessus à l'échelle des provinces et des territoires par groupe d'âge, sexe et facteur de stratification de l'équité.
5. Calcule les taux stratifiés bruts et suit la procédure suivante pour calculer les taux stratifiés normalisés selon l'âge :
 - a) Calcule les taux propres à chaque groupe d'âge.
 - b) Multiplie les taux des groupes d'âge de la population étudiée par le nombre de personnes dans chaque groupe d'âge de la population de référence pour obtenir le taux pondéré selon l'âge de chaque groupe d'âge.
 - c) Pour chaque catégorie de facteur de stratification de l'équité, fait la somme de tous les taux pondérés selon l'âge pour obtenir le taux normalisé selon l'âge.
 - d) Calcule la variance et utilise cette valeur pour calculer les intervalles de confiance.
6. Calcule les taux globaux bruts et normalisés selon l'âge. Notez que les taux globaux correspondent à l'ensemble des catégories d'un facteur de stratification de l'équité (p. ex. pour le facteur de stratification du revenu, le taux global correspond aux quintiles 1 à 5). Ainsi, les cas qui ne sont pas associés à une catégorie de facteur de stratification de l'équité (p. ex. À cause d'un code postal manquant) seront exclus du taux global.
7. Produit le tableau de données comprenant les taux globaux et les taux par facteur de stratification de l'équité.

Code de la macro

%MACRO Calculate_Stratified_Rates(indicator= , yr= , reporting_level= , equity_stratifier= ,
inlib=, infile=, stdpop=, popfile=, popvar=, age_var=, agegpl=, agegpu=, multiplier=, roundunit=,
outlib=, outfile=);

```
/******  
/* STEP 1: Prepare data set of indicator cases aggregated by reporting level, sex, age group  
/* and equity stratifier */  
/******  
/******  
/* 1A. Create count variable for the indicator data set and output required year and age groups  
/******  
data &indicator.&yr._&AGEGPL._&AGEGPU.;  
    set &inlib.&infile.;  
    * Create a count variable for the indicator;  
    &indicator.=1;  
    * Output required year and age groups only;  
    if year eq &yr. and ("&AGEGPL."<=&AGE_VAR.<="&AGEGPU.") then output;  
run;  
  
/******  
/* 1B. Create aggregated indicator data sets - national and by province/territory */  
/* Prepare aggregated counts from indicator data set (infile) by age group, sex  
/* and equity stratifier.  
/* Do this for Canada and by reporting level.  
/******  
* Sum indicator cases by sex, age group and equity stratifier for all of Canada;  
proc sql;  
    create table &indicator._canada as  
    select sex, &AGE_VAR., &equity_stratifier., SUM(&indicator.) as cases  
    from &indicator.&yr._&AGEGPL._&AGEGPU.  
    group by sex, &AGE_VAR., &equity_stratifier.  
    OUTER UNION CORR  
    select "3" AS sex, &AGE_VAR., &equity_stratifier., SUM(&indicator.) as cases  
    from &indicator.&yr._&AGEGPL._&AGEGPU.  
    group by &AGE_VAR., &equity_stratifier.  
    ORDER BY SEX, &AGE_VAR., &equity_stratifier.;  
QUIT;  
  
/* Define Canada as '99' */  
data &indicator._canada;  
    set &indicator._canada;
```

```

        &reporting_level.="99";
run;

* Sum indicator cases by sex, age group, equity stratifier and reporting level;
proc sql;
    create table &indicator._&reporting_level. as
    select &reporting_level., sex, &AGE_VAR., &equity_stratifier., SUM(&indicator.)
    as cases
    from &indicator.&yr._&AGEGPL._&AGEGPU.
    group by &reporting_level., sex, &AGE_VAR., &equity_stratifier.
    OUTER UNION CORR
    select &reporting_level., "3" AS sex, &AGE_VAR., &equity_stratifier., SUM(&indicator.)
    as cases
    from &indicator.&yr._&AGEGPL._&AGEGPU.
    group by &reporting_level., &AGE_VAR., &equity_stratifier.
    ORDER BY &reporting_level., SEX, &AGE_VAR., &equity_stratifier.;
QUIT;

data &indicator._&reporting_level.;
    set &indicator._&reporting_level. &indicator._canada;
run;

/*****
/* STEP 2. Prepare standard population file */
/*****
/*****
/* 2A. Sum the standard population by indicator age groups within age range of the indicator
/* and keep the total population
/* Multiply each of the age-specific rates by the proportion of the 2011 population belonging
/* to the particular age group (called the standard population weight)
/*****
proc sql;
    create table stdpop_agegroup as
    select &AGE_VAR., standard_pop_cnt as stdpop, sum(standard_pop_cnt) as
    agegroup_stdpop
    from &stdpop.(where=("&AGEGPL."<=&AGE_VAR.<="&AGEGPU."));
quit;

proc sort data=stdpop_agegroup; by &AGE_VAR.; run;

proc sql;
    create table stdpop_population as
    select &AGE_VAR., sum(standard_pop_cnt) as total_stdpop
    from &stdpop.(where=("&AGEGPL."<=&AGE_VAR.<="&AGEGPU."));

```

```
quit;

proc sort data=stdpop_population; by &AGE_VAR.; run;

data stdpop_new;
  merge stdpop_agegroup (in=a) stdpop_population(in=b);
  by &AGE_VAR.;
  if a=1 then do;
    weight= stdpop/agegroup_stdpop;
    output;
  end;
  drop total_stdpop;
run;

/*****/
/* STEP 3. Prepare population data */
/*****/
/*****/
/* 3A. Prepare population denominators for the specific age range of indicator */
/*****/
data new_popfile;
  set &popfile.;
  if population_year eq &yr. and ("&AGEGPL."<=&AGE_VAR.<="&AGEGPU.") then output;
run;

/*****/
/* STEP 4. Join data sets from steps 1 to 3 to create datasets for calculating rates by equity
/* stratifier in Step 5 and overall rates in Step 6.
/*****/
/*****/
/* 4A. Create dataset containing population denominator and standard population datasets
/*****/
proc sql;
  create table new_popfile2 as
  select a.*, b.stdpop, b.weight
  from new_popfile as a
  right join stdpop_new as b
  on a.&AGE_VAR.=b.&AGE_VAR.
  order by &reporting_level., a.sex, b.&AGE_VAR., a.&equity_stratifier.;
quit;

/*****/
```

```
/* 4B. Join in dataset of aggregated indicator cases and roll up by equity stratifier,  
/* reporting_level, sex and age.  
/* This will be used for calculating rates by equity stratifier in Step 5.  
/*****/  
proc sql;  
    create table count_pop_&equity_stratifier. as  
    select a.cases as count, b.*  
    from &indicator._&reporting_level. as a  
    right join new_popfile2 as b  
    on a.&AGE_VAR.=b.&AGE_VAR. and a.sex=b.sex and  
    a.&equity_stratifier.= b.&equity_stratifier. and a.&reporting_level.=b.&reporting_level.  
    order by &reporting_level., &equity_stratifier., sex, &AGE_VAR.;
```

quit;

```
*****/  
/* 4C. Create new population file rolling up all &equity_stratifier categories  
*****/  
proc sql;  
    create table new_popfile3 as  
    select population_year, &reporting_level., sex, &AGE_VAR., sum(&popvar.) as  
    &popvar., stdpop, weight  
    from new_popfile2  
    group by population_year, &reporting_level., sex, &AGE_VAR., stdpop, weight;
```

quit;

```
*****/  
/* 4D. Aggregate indicator cases by reporting level, sex and age only  
*****/  
proc sql;  
    create table &indicator._&reporting_level._all  
    as select &reporting_level., sex, &AGE_VAR., sum(cases) as cases  
    from &indicator._&reporting_level.  
    group by &reporting_level., sex, &AGE_VAR.;
```

quit;

```
*****/  
/* 4E. Join datasets from step 4C and 4D.  
    This will be used for calculating overall rates in Step 6.  
*****/  
proc sql;  
    create table count_pop as  
    select a.cases as count, b.*  
    from &indicator._&reporting_level._all as a  
    right join new_popfile3 as b
```

```

        on a.&AGE_VAR.=b.&AGE_VAR. and a.sex=b.sex and
        a.&reporting_level.=b.&reporting_level.
        order by &reporting_level., sex, &AGE_VAR.;
quit;

/*****/
/* STEP 5. Calculate crude and age-standardized rates stratified by equity stratifier
/*****/
/*****/
/* 5A. Calculate crude, expected, variance and standardized numerator count
/* Variance calculation is based on a binomial distribution
/* Standardized numerator count (std_num) is required for Calculate_Inequality_Measures */
/*****/
%if &popvar. ne 0 %then %do;
    data &indicator._rates_&YR._&AGEGPL._&AGEGPU.;
        set count_pop_&equity_stratifier.;
        if count=. then count=0;
        crude=(count/&popvar.)*&MULTIPLIER.;
        expect=crude*weight;
        variance=((weight*weight)*crude*(&MULTIPLIER.-crude))/&popvar.;
        std_num=count/&popvar.*stdpop;
    run;
%end;
%if &popvar. = 0 %then %do;
    data &indicator._rates_&YR._&AGEGPL._&AGEGPU.;
        set count_pop_&equity_stratifier.;
        if count=. then count=0;
        crude=0;
        expect=crude*weight;
        variance=((weight*weight)*crude*(&MULTIPLIER.-crude))/&popvar.;
        std_num=count/&popvar.*stdpop;
    run;
%end;

/*****/
/* 5B. Calculate age-standardized rates (ASR) by equity stratifier categories
/* Sum counts, total population, crude rate, standard numerator, expected and variance
/* by reporting level, sex and equity stratifier
/*****/
proc sql;
    create table &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier. as
    select &reporting_level., sex, &equity_stratifier.,
    sum(count) as count, sum(&popvar.) as totalpop,
    sum(count)/sum(&popvar.)*&MULTIPLIER. as crude, sum(std_num) as std_numer,

```

```

sum(expect) as sum_exp, sum(variance) as variance
from &indicator._rates_&YR._&AGEGPL._&AGEGPU.
group by &reporting_level., sex, &equity_stratifier.;
quit;

/*****/
/* 5C. Calculate confidence intervals for age-standardized rates by equity stratifier
*****/
data &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier.;
  set &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier.;
  lci=sum_exp-1.96*sqrt(variance);
  uci=sum_exp+1.96*sqrt(variance);
  if lci<0 then do;
    p=sum_exp/&MULTIPLIER.;
    if p=0 then p=1/(&MULTIPLIER.*&MULTIPLIER.);
    logitp=log(p/(1-p));
    var=1/(p*p*(1-p)*(1-p))*(variance/(&MULTIPLIER.*&MULTIPLIER.));
    A=logitp-1.96*sqrt(var);
    B=logitp+1.96*sqrt(var);
    lci=&MULTIPLIER./(1+exp(-A));
    uci=&MULTIPLIER./(1+exp(-B));
    drop p logitp var A B;
  end;
  %if &ROUNDUNIT. = 1 %then format crude rstd rlci ruci 8.;
  %else format asr asr_lci asr_uci 8&ROUNDUNIT.;;
  crude_rate= round(crude,&ROUNDUNIT.);
  asr = round(sum_exp,&ROUNDUNIT.);
  asr_lci = round(lci,&ROUNDUNIT.);
  asr_uci = round(uci,&ROUNDUNIT.);
run;

/*****/
/* STEP 6. Calculate overall rates */
/*****/
/*****/
/* 6A. Calculate crude, expected, variance and standardized numerator count
/* Variance calculation is based on a binomial distribution
/* Standardized numerator count (std_num) required in Calculate_Inequality_Measures*/
*****/
%if &popvar. ne 0 %then %do;
  data &indicator._rates_&YR._&AGEGPL._&AGEGPU._ALL;
    set count_pop;
    if count=. then count=0;
    crude=(count/&popvar.)*&MULTIPLIER.;

```

```

        expect=crude*weight;
        variance=((weight*weight)*crude*(&MULTIPLIER.-crude))/&popvar.;
        std_num=count/&popvar.*stdpop;
    run;
%end;
%if &popvar. = 0 %then %do;
    data &indicator._rates_&YR._&AGEGPL._&AGEGPU._ALL;
        set count_pop;
        if count=. then count=0;
        crude=0;
        expect=crude*weight;
        variance=((weight*weight)*crude*(&MULTIPLIER.-crude))/&popvar.;
        std_num=count/&popvar.*stdpop;
    run;
%end;

/*****/
/* 6B. Calculate overall age-standardized rates
/* Sum counts, total population, crude rate, standard numerator, expected and
/* variance by reporting level and sex
/*****/
proc sql;
    create table &indicator._ASR_&YR._&AGEGPL._&AGEGPU. as
    select &reporting_level., sex,
        sum(count) as count, sum(&popvar.) as totalpop,
        sum(count)/sum(&popvar.)*&MULTIPLIER. as crude, sum(std_num) as std_numer,
        sum(expect) as sum_exp, sum(variance) as variance
    from &indicator._rates_&YR._&AGEGPL._&AGEGPU._ALL
    group by &reporting_level., sex;
quit;

/*****/
/* 6C. Calculate confidence intervals for overall age-standardized rates
/*****/
data &indicator._ASR_&YR._&AGEGPL._&AGEGPU.;
    set &indicator._ASR_&YR._&AGEGPL._&AGEGPU.;
    lci=sum_exp-1.96*sqrt(variance);
    uci=sum_exp+1.96*sqrt(variance);
    if lci<0 then do;
        p=sum_exp/&MULTIPLIER.;
        if p=0 then p=1/(&MULTIPLIER.*&MULTIPLIER.);
        logitp=log(p/(1-p));
        var=1/(p*p*(1-p)*(1-p))*(variance/(&MULTIPLIER.*&MULTIPLIER.));
        A=logitp-1.96*sqrt(var);
    end;
end;

```

```

        B=logitp+1.96*sqrt(var);
        lci=&MULTIPLIER./(1+exp(-A));
        uci=&MULTIPLIER./(1+exp(-B));
        drop p logitp var A B;
    end;

    %if &ROUNDUNIT. = 1 %then
        format crude rstd rlci ruci 8.;
    %else
        format asr asr_lci asr_uci 8&ROUNDUNIT.;;
        crude_rate= round(crude,&ROUNDUNIT.);
        asr = round(sum_exp,&ROUNDUNIT.);
        asr_lci = round(lci,&ROUNDUNIT.);
        asr_uci = round(uci,&ROUNDUNIT.);
run;

/*****/
/* STEP 7. Output data set
/*****/
* ASR by equity stratifier;
data &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier.;
    retain year &reporting_level. sex &equity_stratifier count totalpop crude_rate asr
    asr_lci asr_uci variance std_numer;
    set &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier.;
    year=&yr.;
    keep year &reporting_level. sex &equity_stratifier count totalpop crude_rate asr
    asr_lci asr_uci variance std_numer;
run;

* ASR overall;
data &indicator._ASR_&YR._&AGEGPL._&AGEGPU.;
    retain year &reporting_level. sex &equity_stratifier count totalpop crude_rate asr
    asr_lci asr_uci variance std_numer;
    set &indicator._ASR_&YR._&AGEGPL._&AGEGPU.;
    year=&yr.;
    &equity_stratifier ="0";
    keep year &reporting_level. sex &equity_stratifier count totalpop crude_rate asr
    asr_lci asr_uci variance std_numer;
run;

* Combine ASR overall and by equity stratifier;
data &indicator._&YR._&AGEGPL._&AGEGPU._&equity_stratifier._rates;
    set &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier.
    &indicator._ASR_&YR._&AGEGPL._&AGEGPU.;

```

```
label year = 'Year'
      &reporting_level.= 'Reporting level'
      &equity_stratifier. = "&equity_stratifier. "
      sex= 'Sex'
      count = 'Number of indicator cases'
      totalpop = "Total population"
      crude_rate = "Crude rate per &MULTIPLIER. population"
      variance = "Variance"
      asr = "Age-standardized rate per &MULTIPLIER. population"
      asr_lci = "Lower confidence limit of the age-standardized rate"
      asr_uci = "Upper confidence limit of the age-standardized rate"
      std_numer = "Standard numerator";
      if &equity_stratifier. eq . then delete;

run;

proc sort data=&indicator._&YR._&AGEGPL._&AGEGPU._&equity_stratifier._rates;
      by &reporting_level. sex &equity_stratifier.;
run;

/*****
/* Note:      This data set is formatted to run in the Calculate_Inequality_Measures macro.
*****/

data &outlib..&outfile.;
      set &indicator._&YR._&AGEGPL._&AGEGPU._&equity_stratifier._rates;
run;

%MEND Calculate_Stratified_Rates;
```

Annexe C : Macro SAS — %Calculate_Inequality_Measures

Fichiers de sortie

3 fichiers de sortie seront créés après l'exécution de ce programme macro :

- `rd_rr_&indicator.&yr_by_&equity_stratifier.` — ratio des taux et différence des taux selon le facteur de stratification de l'équité et le niveau de déclaration
- `prp_pin_&indicator.&yr_by_&equity_stratifier.` — RPT et NPRP selon le facteur de stratification de l'équité et le niveau de déclaration
- `&outlib.&outfile.` — ratio des taux, différence des taux, RPT et NPRP définitifs selon le facteur de stratification de l'équité et le niveau de déclaration

Paramètres

Vous devez définir tous les paramètres ci-dessous, sauf indication contraire.

- *data* — Nom du fichier de données d'entrée de la macro.
- *indicator* — Nom de l'indicateur de santé mesuré.
- *yr* — Année.
- *equity_stratifier* — Variable pour laquelle vous désirez calculer les mesures de l'inégalité (p. ex. le quintile de revenu).
- *ordered_data* — Paramètre qui précise si votre facteur de stratification de l'équité représente des données ordonnées (p. ex. 5 quintiles de revenu, où le premier quintile est le quintile de revenu le plus faible et le cinquième quintile, le quintile de revenu le plus élevé) ou non ordonnées (p. ex. la langue). La macro accepte 2 valeurs : « yes » (oui) pour les données ordonnées et « no » (non) pour les données non ordonnées.
- *ordered_ref_group* — Pour les facteurs de stratification de l'équité ordonnés, ce paramètre précise si le groupe de référence représente le groupe le plus élevé ou le plus faible. La macro accepte 2 valeurs : « highest » si le groupe de référence représente le groupe le plus élevé ou « lowest » s'il représente le groupe le plus faible. Si vous utilisez des facteurs de stratification de l'équité non ordonnés, laissez ce paramètre vide.
- *unordered_ref_group* — Pour les facteurs de stratification de l'équité non ordonnés, vous devez préciser le groupe de référence; il s'agit d'une valeur représentant le groupe avec lequel tous les autres groupes sont comparés. Si vous utilisez des facteurs de stratification de l'équité ordonnés, laissez ce paramètre vide.
- *reporting_level* — Niveau de déclaration géographique.

- *sex* — Variable du sexe pour l'indicateur; obligatoire même si le rapport concerne les 2 sexes (dans ce cas, entrez une valeur factice).
- *multiplier* — Valeur pour laquelle vous voulez présenter les taux (p. ex. MULTIPLIER = 100,000 pour les taux par 100 000).
- *outlib* — Nom du répertoire SAS où vous voulez stocker votre fichier de sortie.
- *outfile* — Nom du fichier de sortie.

Invocation de la macro

%MACRO Calculate_Inequality_Measures(data=, indicator=, yr=, equity_stratifier=, ordered_data=, ordered_ref_group=, unordered_ref_group=, reporting_level=, sex=, outlib=, outfile=, multiplier=);

L'exemple ci-dessous montre comment la macro peut être invoquée si l'on utilise le fichier de sortie généré après l'exécution de la macro %Calculate_Stratified_Rates à partir des données fictives fournies dans le fichier Excel. Dans cet exemple, on suppose que tous les fichiers sont conservés dans le répertoire *work*.

```
%Calculate_Inequality_Measures(  
    data=work.fakedata_output1,  
    indicator=fake,  
    yr=2020,  
    equity_stratifier=income_quintile,  
    ordered_data=yes,  
    ordered_ref_group=highest,  
    unordered_ref_group=,  
    reporting_level=PT_code,  
    sex=sex,  
    outlib=work,  
    outfile=fakedata_output2,  
    multiplier=100000);
```

Étapes de la macro

1. Crée un fichier de données indiquant le groupe de référence pour chaque combinaison de niveau de déclaration et de sexe.
2. Calcule le ratio des taux et la différence des taux.
3. Calcule la RPT et le NPRP.
4. Crée des tableaux de données comprenant toutes les mesures sommaires de l'inégalité.

Code de la macro

```
%MACRO Calculate_Inequality_Measures(data=, indicator=, yr=, equity_stratifier=,
  ordered_data=, ordered_ref_group=, unordered_ref_group=,
  reporting_level=, sex=, outlib=, outfile=, multiplier=);

/*****/
/* STEP 1: For both ordered and unordered data, create a dataset containing
/* the reference group for each combination of reporting level and sex
*/
/*****/
proc sort data=&data. out=&indicator._sorted;
  by &reporting_level. &sex. &equity_stratifier.;
  where &equity_stratifier. not in ('0', '', '!');
run;

* Case 1: Ordered data and the reference group is the highest group;
%if &ordered_data. eq yes and &ordered_ref_group.=highest %then %do;
data &indicator.&yr._ref;
  set &indicator._sorted;
  by &reporting_level. &sex. &equity_stratifier.;
  rename asr=asr_ref variance=variance_ref &equity_stratifier.=ref;
  if last.&sex. then output;
  keep &reporting_level. &sex. &equity_stratifier. asr variance;
run;
%end;

* Case 2: Ordered data and the reference group is the lowest group;
%if &ordered_data. eq yes and &ordered_ref_group.=lowest %then %do;
data &indicator.&yr._ref;
  set &indicator._sorted;
  by &reporting_level. &sex. &equity_stratifier.;
  rename asr=asr_ref variance=variance_ref &equity_stratifier.=ref;
  if first.&sex. then output;
  keep &reporting_level. &sex. &equity_stratifier. asr variance;
run;
%end;

/* ERROR CONDITION: Unordered data and user has not specified whether the reference
  group is the 'highest' or 'lowest' value */
%if &ordered_data. eq yes and %length(&ordered_ref_group)=0 %then %do;
  %put 'ERROR: For ordered data, specify whether reference group is the 'highest'
    or 'lowest' group.';
  %ABORT;
```

```
%end;

* Case 3: For unordered data, the user must have specified the reference group;
%if &ordered_data. eq no and unordered_ref_group ne . %then %do;
data &indicator.&yr._ref;
    set &indicator._sorted;
    by &reporting_level. &sex. &equity_stratifier.;
    rename asr=asr_ref variance=variance_ref &equity_stratifier.=ref;
    if &equity_stratifier. eq "&unordered_ref_group." then output;
    /* Might need to keep other outputs as well */
    keep &reporting_level. &sex. &equity_stratifier. asr variance;
run;
%end;

* ERROR CONDITION: If unordered data and user has not specified the reference group;
%if &ordered_data. eq no and %length(&unordered_ref_group)=0 %then %do;
    %put 'ERROR: For unordered data, specify a character value for the reference group.';
    %ABORT;
%end;

data &indicator.&yr._ref;
    set &indicator.&yr._ref;
    label asr_ref = "Age standardized rate (Reference group)"
    variance_ref = "Variance (Reference group)"
    ref ="Reference group";
run;

/*****/
/* STEP 2: Calculate rate ratio and rate difference and their confidence limits
/* for stratified data
/*****/
proc sort data=&data. out=&indicator._sorted2;
    by &reporting_level. &sex. &equity_stratifier.;
    where &equity_stratifier. not in ('','!');
run;

/* Link dataset with reference group values to dataset with age standardized rates
for all equity stratifier groups */
proc sql;
    create table &indicator._sorted3
    as select a.*, b.asr_ref as asr_ref, b.variance_ref as variance_ref, b.ref as ref
    from &indicator._sorted2 as a
    left join &indicator.&yr._ref as b
    on a.&reporting_level. = b.&reporting_level. and a.&sex. = b.&sex.;
```

quit;

/* Calculate rate ratio and rate difference and confidence limits */

data &outlib..rd_rr_&indicator._&yr._by_&equity_stratifier.;

set &indicator._sorted3;

/* Rate ratio calculations */

if asr_ref ne . then rr=asr/asr_ref;

else if asr_ref eq . then rr=.;

if rr ne 0 and rr ne . then do;

*** Variance of the log of the rate ratio;**

 var_logrr=((variance/(asr**2))+(variance_ref/(asr_ref**2)));

*** Upper and lower confidence limits of the rate ratio;**

 lcl_rr = exp(log(rr)-(1.96*sqrt(var_logrr)));

 ucl_rr = exp(log(rr)+(1.96*sqrt(var_logrr)));

end;

if rr eq 0 or rr eq . then do;

 var_logrr=.;

*** Upper and lower confidence limits of the rate ratio;**

 lcl_rr = .;

 ucl_rr = .;

end;

/* Rate difference calculations */

if asr_ref ne . then rd=asr-asr_ref;

else if asr_ref eq . then rd=.;

if rd ne . then do;

*** Variance of the rate difference;**

 var_rd=variance + variance_ref;

*** Upper and lower confidence limits of the rate difference;**

 lcl_rd=rd-(1.96*sqrt(var_rd));

 ucl_rd=rd+(1.96*sqrt(var_rd));

end;

if rd eq . then do;

*** Variance of the rate difference;**

 var_rd=.;

*** Upper and lower confidence limits of the rate difference;**

 lcl_rd=.;

 ucl_rd=.;

end;

```

        /* If reporting for the reference group the variance and CIs should not be reported */
        if &equity_stratifier. eq ref then do;
            var_logrr=.;
            lcl_rr = .;
            ucl_rr = .;
            var_rd=.;
            lcl_rd=.;
            ucl_rd=.;
        end;
run;

/* We do not report the rate difference and the rate ratio for the overall category ('0')
but need to keep the overall category to join to PRR and PIN results in Step 4*/
data &outlib..rd_rr_&indicator._&yr._by_&equity_stratifier.;
set &outlib..rd_rr_&indicator._&yr._by_&equity_stratifier.;
if &equity_stratifier. eq '0' then do;
    rr=.;
    lcl_rr = .;
    ucl_rr = .;
    var_logrr=.;
    rd=.;
    lcl_rd=.;
    ucl_rd=.;
    var_rd=.;
end;
run;

/*****
/* STEP 3: Calculate PRR and PIN */
*****/

/* 3A. CALCULATE POTENTIAL RATE REDUCTION (PRR)
Requires:
- Pi = proportion of the population in the ith category (by reporting level and sex)
-- Use the total population (totalpop) to calculate the population proportion for each
equity stratifier category by reporting level and sex
- Rate ratios: (rate i/rate reference)
*/
*****/
* Calculate population proportion for each equity stratifier category by reporting level
and sex;
proc sql;
create table prr_&indicator.&yr._a as
select &reporting_level., &sex., &equity_stratifier., totalpop, asr, variance,

```

```
(totalpop/sum(totalpop)) as pop_portion
from &indicator._sorted
group by &reporting_level., &sex.;
```

quit;

* Create dataset containing population proportions and calculated rate ratios;

proc sql;

```
create table prr_&indicator.&yr._b as
select a.&reporting_level., a.&sex., a.&equity_stratifier., a.totalpop, a.asr, a.variance,
a.pop_portion, b.ref as ref, b.asr_ref as asr_ref, a.asr/asr_ref as rate_ratio
from prr_&indicator.&yr._a as a inner join &indicator.&yr._ref as b
on a.&reporting_level. = b.&reporting_level. and a.&sex. = b.&sex.;
```

quit;

* Calculate PRR by reporting level and sex;

proc sql;

```
create table prr_&indicator.&yr. as
select &reporting_level., &sex.,
sum(pop_portion*(rate_ratio-1)) / (1+sum(pop_portion*(rate_ratio-1))) as prr
from prr_&indicator.&yr._b
group &reporting_level., &sex.;
```

quit;

/*****/

/* 3B. CALCULATE CONFIDENCE LIMITS FOR PRR

Require:

- Variance(rate ratio^{1:4,5}) – for formulas see SAS macros and methodology notes

*/

/*****/

* i) Get values for comparison groups to calculate $\text{Sum}(\text{Pi} \cdot \text{Pi} \cdot \text{variance}(\text{rate}))$

AND also $\text{Pi} \cdot \text{rate}$ (values for comparison groups are needed in subsequent calculations for CIs);

proc sql;

```
Create table rate_comp_&indicator.&yr. as
select &reporting_level., &sex., sum((pop_portion**2)*variance) as variance_comp,
sum(pop_portion*asr) as sum_comp
from prr_&indicator.&yr._b as a
where &equity_stratifier. NE ref
group by &reporting_level., &sex.
order by &reporting_level., &sex., &equity_stratifier.;
```

quit;

* ii) Get values for reference group to be used in subsequent CI calculations;

```
proc sql;
    create table rate_ref_&indicator.&yr. as
    select &reporting_level., &sex., variance as variance_ref, asr as rate_ref,
    pop_portion as pop_portion_ref
    from prr_&indicator.&yr._b as a
    where &equity_stratifier. EQ ref
    order by &reporting_level., &sex., &equity_stratifier.;
```

quit;

* iii) Calculate upper and lower CIs for the PRR using the datasets created in step i and ii;

```
proc sql;
    create table prr_ci_&indicator.&yr. as
    select a.&reporting_level., a.&sex., prr,
    (sum_comp/rate_ref) as rate_comp,
    ((variance_comp/sum_comp**2)+(variance_ref/rate_ref**2)) as varlog,
    exp(log(calculated rate_comp)-1.96*sqrt(calculated varlog)) as lcl_low,
    exp(log(calculated rate_comp)+1.96*sqrt(calculated varlog)) as ucl_low,
    1-(1/(pop_portion_ref+calculated lcl_low)) as lcl_prr,
    1-(1/(pop_portion_ref+calculated ucl_low)) as ucl_prr
    from prr_&indicator.&yr. as a, rate_comp_&indicator.&yr. as b,
    rate_ref_&indicator.&yr. as c
    where a.&reporting_level.=b.&reporting_level.=c.&reporting_level.
    and a.&sex.=b.&sex.=c.&sex.;
```

quit;

```
/******  

/* 3C. CREATE DATASET WITH PRR, confidence intervals and  

    Population Impact Number (PIN) */  

/******
```

```
proc sql;
    create table prr_pin_&indicator.&yr._by_&equity_stratifier. as
    select &yr. as year, a.&reporting_level., a.&sex., a.&equity_stratifier.,
    prr, lcl_prr, ucl_prr, std_numer*prr as pin
    from &data. as a left join prr_ci_&indicator.&yr. as b
    on a.&reporting_level.=b.&reporting_level. and a.&sex.=b.&sex.;
```

quit;

* Report PRR and associated confidence intervals as a percentage;

```
data prr_pin_&indicator.&yr._by_&equity_stratifier.;
    set prr_pin_&indicator.&yr._by_&equity_stratifier.;
    prr=prr*100;
    lcl_prr=lcl_prr*100;
    ucl_prr=ucl_prr*100;
```

run;

```

proc sort data=pr_r_pin_&indicator.&yr._by_&equity_stratifier.
    out=&outlib..pr_r_pin_&indicator.&yr._by_&equity_stratifier. nodup;
    by &reporting_level. &sex. &equity_stratifier.;
run;

/* We only report the PRR and PIN for the overall category ('0') but need to
keep the other categories to join to RD and RR results in Step 4*/
data &outlib..pr_r_pin_&indicator.&yr._by_&equity_stratifier.;
    set &outlib..pr_r_pin_&indicator.&yr._by_&equity_stratifier.;
    if &equity_stratifier. ne '0' then do;
        pr_r=.;
        lcl_pr_r=.;
        ucl_pr_r=.;
        pin=.;
    end;
run;

/*****
/* STEP 4: Output final data set with all rates and summary measures of inequality
*/
*****/
proc sql;
    create table &indicator._&equity_stratifier._inequalities as
    select a.year, a.&reporting_level., a.&equity_stratifier., a.&sex., a.count, a.totalpop,
    a.crude_rate as crude_rate, a.asr, a.asr_lci, a.asr_uci , a.rr, a.lcl_rr, a.ucl_rr, a.rd,
    a.lcl_rd, a.ucl_rd, b.prr, b.lcl_prr, b.ucl_prr, b.pin
    from &outlib..rd_rr_&indicator._&yr._by_&equity_stratifier. as a,
    &outlib..pr_r_pin_&indicator.&yr._by_&equity_stratifier. as b
    where a.&reporting_level.=b.&reporting_level. and a.&sex.=b.&sex.
    and a.&equity_stratifier. = b.&equity_stratifier.
    order by &reporting_level., &sex., &equity_stratifier.;
quit;

data &outlib..&outfile.;
    retain year &reporting_level. &equity_stratifier &sex. count totalpop crude_rate asr
    asr_lci asr_uci rr lcl_rr ucl_rr rd      lcl_rd ucl_rd pr_r lcl_prr ucl_prr pin;
    length &equity_stratifier. $12.;
    set &indicator._&equity_stratifier._inequalities;

    label year = 'Year'
           &reporting_level. = 'Reporting level'
           &equity_stratifier. = "&equity_stratifier. "

```

```
&sex.          = 'Sex'  
              Count = 'Number of indicator cases'  
              totalpop = 'Total population'  
              crude_rate = "Crude rate per &MULTIPLIER. population"  
              asr = "Age-standardized rate per &MULTIPLIER. population"  
              asr_lci = 'Lower confidence limit of the age-standardized rate'  
              asr_uci = 'Upper confidence limit of the age-standardized rate'  
              rr = 'Rate Ratio (RR)'  
              lcl_rr = 'Lower confidence limit of RR'  
              ucl_rr = 'Upper confidence limit of RR'  
              rd = "Rate Difference (RD) per &MULTIPLIER. population"  
              lcl_rd = 'Lower confidence limit of RD'  
              ucl_rd = 'Upper confidence limit of RD'  
              prr = 'Potential Rate Reduction (PRR) (%)'  
              lcl_prr = 'Lower confidence limit of PRR (%)'  
              ucl_prr = 'Upper confidence limit of PRR (%)'  
              pin = "Population Impact Number (PIN) per &MULTIPLIER. population";  
              keep year &reporting_level. &equity_stratifier. &sex. count totalpop crude_rate asr  
              asr_lci asr_uci rr lcl_rr ucl_rr rd lcl_rd ucl_rd prr lcl_prr ucl_prr pin;  
  
run;  
  
%mend;
```

Annexe D : Fonction R — Calculate_Stratified_Rates

Fichier de sortie

Cette fonction produit un fichier de données contenant des taux bruts et normalisés selon l'âge pour le facteur de stratification de l'équité et le niveau de déclaration choisis.

Paramètres

Vous devez définir tous les paramètres ci-dessous, à l'exception des paramètres comportant des valeurs par défaut, lesquels peuvent être définis au besoin.

- yr — Année. Les données sur les cas liés à l'indicateur et les estimations de la population seront uniquement extraites pour l'année indiquée. Une seule année peut être saisie à la fois.
- reporting_level — Niveau de déclaration géographique. La fonction calcule les taux à l'échelle du Canada et des provinces et territoires. Donnez à la variable reporting_level le même nom dans votre fichier des cas liés à l'indicateur (infile) et votre fichier des estimations de la population (popfile).
- equity_stratifier — Variable pour laquelle vous désirez calculer les taux stratifiés (p. ex. le quintile de revenu). Il peut s'agir d'un facteur de stratification ordonné, tel que le quintile de revenu, ou d'un facteur de stratification non ordonné, tel que l'emplacement géographique (région urbaine ou région rurale/éloignée). Vous devez utiliser le même nom pour la variable equity_stratifier dans tous vos fichiers d'entrée.
- infile — Nom du fichier d'entrée des cas liés à l'indicateur.
- stdpop — Nom du fichier des estimations de la population de référence.
- popfile — Nom du fichier des estimations de la population catégorisées selon le facteur de stratification de l'équité choisi (p. ex. estimations de la population par quintile de revenu).
- popvar — Chiffres concernant la population provenant du fichier des estimations de la population (popfile).
- age_var — Nom de la variable de l'âge utilisée uniformément dans le fichier des cas liés à l'indicateur (infile), le fichier des estimations de la population de référence (stdpop) et le fichier des estimations de la population (popfile).
- agegl — Limite d'âge inférieure de l'indicateur de santé; doit respecter le format 1, 2, [...], 14, 15 tel que précisé dans la variable de l'âge (p. ex. si la limite d'âge inférieure pour votre indicateur de santé est 0 an, cette valeur doit être 1). La valeur par défaut est 1.
- agegpu — Limite d'âge supérieure de l'indicateur de santé; doit respecter le format 1, 2, [...], 18, 19 tel que précisé dans la variable de l'âge (p. ex. si la limite d'âge supérieure pour votre indicateur de santé est 74 ans, cette valeur doit être 15). La valeur par défaut est 19.

- multiplier — Valeur pour laquelle vous voulez présenter les taux (p. ex. multiplier = 100,000 pour les taux par 100 000). La valeur par défaut est 100 000.
- roundunit — Valeur d'arrondissement (p. ex. pour arrondir à 2 décimales, précisez ROUNDUNIT = 2). La valeur par défaut est 2.

Invocation de la fonction

Calculate_Stratified_Rates (yr=, reporting_level=, equity_stratifier=, infile=, stdpop=, popfile=, popvar=, age_var=, agegpu=1, agegpl=19, multiplier=100000, roundunit=2)

L'exemple ci-dessous montre comment la fonction peut être invoquée si l'on utilise les données fictives fournies dans le fichier Excel.

Calculate_Stratified_Rates(yr=2020, reporting_level=PT_code, equity_stratifier=income_quintile, infile=Example_infile_data, stdpop=Example_stdpop_file, popfile=Example_popfile_data, popvar=PT_pop, age_var=AGE_GROUP_CODE, agegpu=1, agegpl=19, multiplier=100000, roundunit=2)

Étapes de la fonction

1. Prépare les dénominateurs de population pour la tranche d'âge et l'année visées.
Crée des chiffres globaux concernant la population pour l'ensemble des niveaux de stratification de l'équité.
2. Prépare les cas liés à l'indicateur en créant des données agrégées pour chaque combinaison de niveau de déclaration, de groupe d'âge, de sexe et de facteur de stratification de l'équité.
Crée d'autres chiffres globaux concernant la population pour les facteurs de stratification de l'équité, les sexes et les niveaux provinciaux et territoriaux (p. ex. crée des chiffres à l'échelle nationale).
3. Prépare le fichier de population de référence qui servira à la normalisation selon la tranche d'âge visée.
4. Fusionne les fichiers des cas liés à l'indicateur et des chiffres concernant la population.
5. Défusionne les fichiers des cas liés à l'indicateur et des chiffres concernant la population, et les transpose de façon à ce que chaque combinaison de sexe, de province ou territoire et de facteur de stratification corresponde à une colonne, et chaque groupe d'âge, à une rangée.

6. Calcule les taux bruts stratifiés et globauxⁱⁱⁱ, et suit la procédure suivante pour calculer les taux normalisés selon l'âge stratifiés et globaux :
 - a. Calcule les taux propres à chaque groupe d'âge.
 - b. Multiplie les taux des groupes d'âge de la population étudiée par le nombre de personnes dans chaque groupe d'âge de la population de référence pour obtenir le taux pondéré selon l'âge de chaque groupe d'âge.
 - c. Pour chaque catégorie de facteur de stratification de l'équité, fait la somme de tous les taux pondérés selon l'âge pour obtenir le taux normalisé selon l'âge.
 - d. Calcule la variance et utilise cette valeur pour calculer les intervalles de confiance.
7. Produit le tableau de données comprenant les taux globaux et les taux par facteur de stratification de l'équité.

Remarques supplémentaires

Ce code **invoque 2** paquetages d'un répertoire externe : `dplyr`³ et `reshape2`⁴.

Code de la fonction

```
"Calculate_Stratified_Rates" <-  
function (reporting_level, age_var, equity_stratifier, infile, popfile,  
          popvar, stdpop, roundunit=2, multiplier=100000,  
          agegpl=1, agegpu=19, yr)  
{  
  
  library(reshape2)  
  library(dplyr)  
  
  ### rename popvar variable to PT_pop  
  popfile <- popfile %>%  
    rename(PT_pop=!!enquo(popvar))  
  
  ### subset popfile and infile so only working with the required columns,  
  # and remove any rows with missing values  
  popfile <- popfile %>%  
    select(POPULATION_YEAR, !!enquo(reporting_level), sex,  
           !!enquo(age_var), !!enquo(equity_stratifier),  
           PT_pop) %>%
```

iii. Les taux globaux correspondent à l'ensemble des catégories d'un facteur de stratification de l'équité (p. ex. pour le facteur de stratification du revenu, le taux global correspond aux quintiles 1 à 5). Ainsi, les cas qui ne sont pas associés à une catégorie de facteur de stratification de l'équité (p. ex. à cause d'un code postal manquant) seront exclus du taux global.

```
na.omit

infile <- infile %>%
  select(YEAR, !!enquo(reporting_level), sex,
  !!enquo(age_var), !!enquo(equity_stratifier)) %>%
  na.omit

### format popfile

#create overall counts by equity stratifier, where equity stratifier=0
popfile_long_ES_agg <- aggregate(x = popfile[c("PT_pop")],
  by = popfile[c(deparse(substitute(reporting_level)),
  "sex", deparse(substitute(age_var)),
  "POPULATION_YEAR")], FUN = sum) %>%
  mutate(!!enquo(equity_stratifier) := "0") %>%
  bind_rows(popfile) %>% #combine with popfile
  mutate(pt_sex_es = paste(!!enquo(reporting_level), sex,
  !!enquo(equity_stratifier), sep = "_")) %>%
  filter (
  as.numeric(POPULATION_YEAR) == yr #only keep data from relevant year
  )

#filter by age parameter
popfile_long_ES_agg <-
  popfile_long_ES_agg[which(as.numeric(popfile_long_ES_agg
  [[deparse(substitute(age_var))]])
  >= agegpu
  & as.numeric(popfile_long_ES_agg
  [[deparse(substitute(age_var))]])
  <= agegpu), ]

#list population counts by pt_sex_es variable and age group
popfile_tab <- aggregate(x = popfile_long_ES_agg[c("PT_pop")],
  by = popfile_long_ES_agg[c("pt_sex_es",
  deparse(substitute(age_var))]),
  FUN = max)

### format infile

infile_long_sub <- infile %>%
  filter (as.numeric(YEAR) == yr) #only keep data from relevant year

#filter by age parameter
infile_long_sub <-
```

```

infile_long_sub [which(as.numeric(infile_long_sub
                        [[deparse(substitute(age_var))]])
                        >= agegpl
                        & as.numeric(infile_long_sub
                        [[deparse(substitute(age_var))]])
                        <= agegpu), ]

#create indicator counts by reporting level, sex, age, and
#equity stratifier
infile_tab <- infile_long_sub %>%
  count (!!enquo(reporting_level), sex, !!enquo(age_var),
        !!enquo(equity_stratifier), YEAR)

#create overall counts by equity stratifier, where equity stratifier=0
infile_tab_ES_agg <-
  aggregate(x = infile_tab[c("n")],
           by = infile_tab
           [c(deparse(substitute(reporting_level)),
             "sex", deparse(substitute(age_var)), "YEAR")],
           FUN = sum) %>%
  mutate (!!enquo(equity_stratifier) := "0") %>%
  bind_rows(infile_tab) #combine with the counts created in step above

#create overall counts by sex, where sex=3
infile_tab_ES_sex_agg <- aggregate(x = infile_tab_ES_agg[c("n")],
  by = infile_tab_ES_agg[c(deparse(substitute(reporting_level)),
    deparse(substitute(equity_stratifier)),
    deparse(substitute(age_var)), "YEAR")],
  FUN = sum) %>%
  mutate(sex = "3") %>%
  bind_rows(infile_tab_ES_agg) #combine with counts created in step above

#create overall counts by reporting level, reporting level=99
infile_tab_ES_sex_PT_agg <-
  aggregate(x = infile_tab_ES_sex_agg[c("n")],
           by = infile_tab_ES_sex_agg[c("sex",
           deparse(substitute(equity_stratifier)),
           deparse(substitute(age_var)), "YEAR")],
           FUN = sum)%>%
  mutate (!!enquo(reporting_level) := "99") %>%
  bind_rows(infile_tab_ES_sex_agg) %>% #combine with counts created above
  mutate(pt_sex_es = paste (!!enquo(reporting_level), sex,
    !!enquo(equity_stratifier), sep="_")) %>%
  select (!!enquo(age_var), pt_sex_es, n)

```

```
#list indicator counts by pt_sex_es variable and age group
```

```
### Merge the indicator and population files, so that they match  
# by pt_sex_es, and age group. Then remove any combinations of PT*sex*ES in  
# indicator cases that do not have a corresponding population count.  
# For combinations of PT*sex*ES in population counts that do not have a  
# corresponding indicator case, set number of indicator cases to zero
```

```
merged_pop_infile <- merge(popfile_tab, infile_tab_ES_sex_PT_agg ,  
                           by.x=c("pt_sex_es",  
                                   deparse(substitute(age_var))),  
                           by.y=c("pt_sex_es",  
                                   deparse(substitute(age_var))),  
                           all.x=TRUE) %>%  
  mutate(n=ifelse(is.na(n), 0, n))
```

```
### Unmerge the indicator and population file at the end,  
# so they can be transformed into format needed for  
# "calc.age.adjusted.rates"
```

```
pop_data <- recast( merged_pop_infile[,c('pt_sex_es',  
                                       deparse(substitute(age_var)), 'PT_pop')],  
                  formula(paste(substitute(age_var), "+",  
                                substitute(variable),  
                                "~", substitute(pt_sex_es))),  
                  id.var = c("pt_sex_es",  
                             deparse(substitute(age_var))))  
#change to a wide format
```

```
#order from lowest to highest age group
```

```
pop_data <- pop_data[order(  
  as.numeric(pop_data[[deparse(substitute(age_var))]]), )
```

```
pop_data[,-c(1,2)][pop_data[,-c(1,2)]==0] <- NA  
#set any 0 population counts to NA so calculations can skip the NA's
```

```
count_data <-  
  recast( merged_pop_infile[,c('pt_sex_es',  
                               deparse(substitute(age_var)), 'n')],  
         formula(paste(substitute(age_var), "+",  
                       substitute(variable), "~",  
                       substitute(pt_sex_es))),
```

```
id.var = c("pt_sex_es", deparse(substitute(age_var)))
#change to a wide format

#order from lowest to highest age group
count_data <-
  count_data[order(as.numeric(
    count_data[[deparse(substitute(age_var))]]), )

#####format stdpop to stdpop_list

#filter by age parameter
stdpop_data_sub <-
  stdpop[which(as.numeric(
    stdpop[[deparse(substitute(age_var))]] >= agegpl
  & as.numeric(
    stdpop[[deparse(substitute(age_var))]] <= agegpu), ]

stdpop_list<- stdpop_data_sub$STANDARD_POP_CNT
names(stdpop_list) <- stdpop_data_sub[[deparse(substitute(age_var))]]
stdpop_list <- stdpop_list[order(as.numeric(names(stdpop_list)))]
#order from lowest to highest age group

###function to create direct age adjusted stratified rates for each
#combo of sex*P/T*Equity stratifier

"calc.age.adjusted.rates" <-
function (count, pop, stdpop)
{
  if(missing(pop))
    pop <- numeric(0)
  rate <- count/pop
  cruderate <- sum(count, na.rm = TRUE)/sum(pop,na.rm = TRUE)
  stdwt <- stdpop/sum(stdpop)
  dsr <- sum(stdwt * rate, na.rm = TRUE)
  variance=sum((stdwt^2)*(count/pop*multiplier)*
    ((multiplier-(count/pop*multiplier))/pop), na.rm=TRUE)
  lci <- dsr*multiplier - 1.96*sqrt(variance)
  uci <- dsr*multiplier + 1.96*sqrt(variance)
  if (lci < 0 ) {
    p <- dsr
    if (p==0) {
      p <- 1/(multiplier^2)}
    logitp <- log(p/(1-p))
```

```
var <- 1/(p*p*(1-p)*(1-p))*(variance/(multiplier^2))
A <- logitp-1.96*sqrt(var)
B <- logitp+1.96*sqrt(var)
lci <- multiplier/(1+exp(-A))
uci <- multiplier/(1+exp(-B))
}

c(year=yr,
  PT_code = strsplit(colnames(count), "_")[[1]][1],
  sex = strsplit(colnames(count), "_")[[1]][2],
  Equity_stratifier = strsplit(colnames(count), "_")[[1]][3],
  number_indicator_cases = sum(count, na.rm=TRUE),
  totalpop = sum(pop, na.rm=TRUE),
  crude_rate = round(cruderate*multiplier, roundunit),
  variance=round(variance, roundunit),
  asr = round(dsr*multiplier, roundunit),
  asr_lci = round(lci, roundunit),
  asr_uci = round(uci, roundunit),
  std_numer = round(sum(count/pop*stdpop, na.rm=TRUE), roundunit))
}
```

###apply the function across the prepared data frames

```
final <- data.frame(matrix(NA, nrow = 0, ncol = 12))

for (i in 3:ncol(count_data)) {
  res <- c(calc.age.adjusted.rates(count_data[i],
                                  pop_data[i],stdpop=stdpop_list))
  final[nrow(final) + 1,] <- res
}

colnames(final) <- c("year", deparse(substitute(reporting_level)),
  "sex", deparse(substitute(equity_stratifier)),
  "count", "totalpop", "crude_rate", "variance",
  "asr", "asr_lci", "asr_uci", "std_numer")
```

set all columns as numeric values, except for stratifier columns

```
final <- as.data.frame(lapply(final,as.numeric))

cols.char <- c(deparse(substitute(reporting_level)),
  "sex", deparse(substitute(equity_stratifier)))
```

```
final[cols.char] <- lapply(final[cols.char],as.character)
```

```
### output data
```

```
final
```

```
}
```

Annexe E : Fonction R — Calculate_Inequality_Measures

Fichier de sortie

Cette fonction produit un fichier de données contenant les résultats finaux relatifs au RT, à la DT, à la RPT et au NPRP selon le facteur de stratification de l'équité et le niveau de déclaration.

Paramètres

Vous devez définir tous les paramètres ci-dessous.

- *data* — Nom du fichier de données d'entrée de la fonction.
- *equity_stratifier* — Variable pour laquelle vous désirez calculer les mesures de l'inégalité (p. ex. le quintile de revenu).
- *ref_group* — Vous devez préciser le groupe de référence. Il s'agit d'une valeur numérique représentant le groupe avec lequel tous les autres groupes sont comparés. Vous ne pouvez pas utiliser 0, car cette valeur est réservée au calcul des taux globaux.
- *reporting_level* — Niveau de déclaration géographique.
- *sex* — Variable du sexe pour l'indicateur; obligatoire même si le rapport concerne les 2 sexes (dans ce cas, entrez une valeur factice).

Invocation de la fonction

Calculate_Inequality_Measures (data=, equity_stratifier=, ref_group=, reporting_level=, sex=)

L'exemple ci-dessous montre comment la fonction peut être invoquée si l'on utilise le fichier de sortie généré après l'exécution de la fonction Calculate_Stratified_Rates à partir des données fictives fournies dans le fichier Excel.

```
Calculate_Inequality_Measures(data=fakedata_output1, equity_stratifier=income_quintile,  
ref_group=5, reporting_level=PT_code, sex=sex)
```

Étapes de la fonction

1. Crée un fichier de données indiquant le groupe de référence pour chaque combinaison de niveau de déclaration et de sexe.
2. Calcule le ratio des taux et la différence des taux.
3. Calcule la RPT et le NPRP.
4. Crée des tableaux de données comprenant toutes les mesures sommaires de l'inégalité.

Remarques supplémentaires

- Ce code **invoque un** paquetage d'un répertoire externe : dplyr³.
- Les taux bruts et normalisés selon l'âge ainsi que la DT et le NPRP présentés dans le fichier de sortie de cette fonction doivent être interprétés en fonction du multiplicateur appliqué au fichier d'entrée (p. ex. si les taux normalisés selon l'âge dans le fichier d'entrée avaient été calculés par 100 000, le NPRP sera également présenté selon un multiplicateur de 100 000).

Code de la fonction

```
"Calculate_Inequality_Measures" <- function (data, equity_stratifier,
                                             ref_group, reporting_level, sex)
{
  library(dplyr)

  ###will produce error is user specifies 0 as the reference group
  if(ref_group == 0) stop(
    'ref_group cannot be specified as 0. Value 0 is reserved for overall rate')

  ###rename parameter variables
  data <- data %>%
    rename(reporting_var=!!enquo(reporting_level)) %>%
    rename(equity_var=!!enquo(equity_stratifier)) %>%
    rename(sex_var = !!enquo(sex))

  ###extract asr and variance of the reference level
  refgroups <- data %>% filter(equity_var==paste(ref_group)) %>%
    mutate (asr_ref = asr, variance_ref = variance,
            ref=ref_group) %>%
    select(year, reporting_var, sex_var, asr_ref,
           variance_ref, ref)

  ###calculate RR and RR confidence intervals
  #(except for when equity stratifier is 0).
  #Don't calculate confidence intervals if rr=0 or NA
  data_rr <- merge(data, refgroups,
                   by=c("sex_var", "reporting_var", "year")) %>%
    mutate(rr = ifelse(equity_var=='0', NA, asr/asr_ref)) %>%
    mutate(var_logrr = ifelse(!is.na(rr) | rr != 0 ,
                              (variance/(asr**2))+ (variance_ref/
                                                    (asr_ref**2)),
                              NA)) %>%
```

```
mutate(lcl_rr = ifelse(!is.na(rr) | rr != 0 ,
                      exp(log(rr)-(1.96*sqrt(var_logrr))),
                      NA)) %>%
mutate(ucl_rr = ifelse(!is.na(rr) | rr != 0 ,
                      exp(log(rr)+(1.96*sqrt(var_logrr))),
                      NA))

###calculate RD and RD confidence intervals
#(except for when equity stratifier is 0).
#Don't calculate confidence intervals if rd=NA
data_rd <- data_rr %>%
  mutate(rd = ifelse(equity_var=='0', NA, asr-asr_ref)) %>%
  mutate(var_rd = ifelse(!is.na(rd),
                        variance+variance_ref,
                        NA)) %>%
  mutate(lcl_rd = ifelse(!is.na(rd),
                        rd-(1.96*sqrt(var_rd)),
                        NA)) %>%
  mutate(ucl_rd = ifelse(!is.na(rd),
                        rd+(1.96*sqrt(var_rd)),
                        NA)) %>%

#set all variance or CI estimates for the reference levels to NA
mutate(
  var_logrr=replace(var_logrr, equity_var == as.character(ref), NA),
  lcl_rr=replace(lcl_rr, equity_var == as.character(ref), NA),
  ucl_rr=replace(ucl_rr, equity_var == as.character(ref), NA),
  var_rd=replace(var_rd, equity_var == as.character(ref), NA),
  lcl_rd=replace(lcl_rd, equity_var == as.character(ref), NA),
  ucl_rd=replace(ucl_rd, equity_var == as.character(ref), NA))

###Calculations for PRR

data_prr <- data_rd %>%
  #Calculate population proportion for each equity stratifier category
  # by reporting level and sex
  filter (equity_var != "0") %>%
  group_by(sex_var, reporting_var) %>%
  mutate(totalpop_sex_var_geo = sum(totalpop, na.rm = TRUE)) %>%
  ungroup() %>%
  mutate(pop_portion = totalpop/totalpop_sex_var_geo) %>%
  group_by(sex_var, reporting_var) %>%
  #calculate prr
  mutate(prr = (sum(pop_portion*(rr-1), na.rm=TRUE) /
```

```

(1+sum(pop_portion*(rr-1),na.rm=TRUE)))) %>%
#copy the prr value into each row of equity stratifier variables
mutate(pop_portion_ref = ifelse(equity_var==as.character(ref),
                               pop_portion, NA) ) %>%
mutate(pop_portion_ref =
        pop_portion_ref[which(equity_var==as.character(ref))])

#get values for comparison groups (need for subsequent calculations)
data_prr_comp <- data_prr %>%
  filter (equity_var!=as.character(ref)) %>%
  group_by(sex_var, reporting_var) %>%
  mutate(variance_comp = sum((pop_portion**2)*variance,
                             na.rm=TRUE),
         sum_comp = sum(pop_portion*asr, na.rm=TRUE))

# calculate prr confidence intervals
data_prr_cl <- data_prr_comp %>%
  mutate(rate_comp = sum_comp/asr_ref,
         varlog = ((variance_comp/sum_comp**2)+(variance_ref/
                                                  asr_ref**2)),
         lcl_low = exp(log(rate_comp)-1.96*sqrt(varlog)),
         ucl_low = exp(log(rate_comp)+1.96*sqrt(varlog)),
         lcl_prr = 1-(1/(pop_portion_ref+lcl_low)),
         ucl_prr = 1-(1/(pop_portion_ref+ucl_low)),
         prr = prr*100,
         lcl_prr = lcl_prr*100,
         ucl_prr = ucl_prr*100) %>%
  select(year, sex_var, reporting_var, equity_var, prr,
         lcl_prr, ucl_prr) %>%
  mutate(equity_var = '0') %>%
  distinct()

#resulting dataset with one row for each reporting level and sex,
# equity_var=0

#merge the rd and rr results with the prr results for final output table
data_output <- merge(data_rd, data_prr_cl,
                    by=c("sex_var", "reporting_var", "year",
                        "equity_var"), all.x = TRUE) %>%
  mutate(pin = prr*std_numer/100) %>% #calculate PIN
  select(year,
         reporting_var, # Reporting level
         equity_var, # Equity stratifier
         sex_var, # Sex
         count, # Number of indicator cases

```

```
totalpop, # Total population
crude_rate, # Crude rate per your
            # multipliers population
asr, # Age-standardized rate per
      # your multipliers population
asr_lci, # Lower confidence limit of the asr
asr_uci, # Upper confidence limit of the asr
rr, # Rate Ratio
lcl_rr, # Lower confidence limit of RR
ucl_rr, # Upper confidence limit of RR
rd, # Rate Difference per your
      # multipliers population
lcl_rd, # Lower confidence limit of RD
ucl_rd, # Upper confidence limit of RD
pr, # Potential Rate Reduction (PRR) (%)
lcl_pr, # Lower confidence limit of PRR (%)
ucl_pr, # Upper confidence limit of PRR (%)
pin) %>% # Population Impact Number (PIN)
          # per your multipliers population
arrange(reporting_var, sex_var, equity_var) %>%
rename(!!enquo(reporting_level):=reporting_var) %>%
rename(!!enquo(equity_stratifier):=equity_var) %>%
rename(!!enquo(sex):=sex_var )
```

data_output

}

Références

1. Rothman K, Greenland S, Lash T. *Modern Epidemiology*. 2008.
2. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *American Journal of Public Health*. Janvier 1998.
3. Wickham H et al. [dplyr: A Grammar of Data Manipulation — R package version 1.0.7](#). 2021.
4. Wickham H. [Reshaping data with the reshape package](#). *Journal of Statistical Software*. 21(12), 1-20. 2007.



ICIS Ottawa

495, chemin Richmond
Bureau 600
Ottawa (Ont.)
K2A 4H6
613-241-7860

ICIS Toronto

4110, rue Yonge
Bureau 300
Toronto (Ont.)
M2P 2B7
416-481-2002

ICIS Victoria

880, rue Douglas
Bureau 600
Victoria (C.-B.)
V8W 2B7
250-220-4100

ICIS Montréal

1010, rue Sherbrooke Ouest
Bureau 602
Montréal (Qc)
H3A 2R7
514-842-2226

icis.ca

10152-0522

