Les problèmes avec le premier chapitre du rapport Fraser : écueils dans l'analyse des tendances temporelles de l'IMC

Ernest Lo, Ph.D.

RÉSUMÉ

Le premier chapitre du rapport *Obesity in Canada: Overstated Problems, Misguided Policy Solutions* de l'Institut Fraser présente une analyse imparfaite et trompeuse des tendances temporelles de l'IMC. Par l'examen de ces imperfections, ce commentaire se veut un tutoriel sur l'analyse des tendances temporelles de l'IMC.

Trois problèmes sont abordés :

1. Repérer les zones de chevauchement des intervalles de confiance est une méthode statistiquement imparfaite pour évaluer des tendances; les méthodes de régression, qui mesurent le comportement des données dans leur ensemble, sont préférables.
2. La stabilité temporelle de la prévalence de l'embonpoint ($25 \leq \text{IMC} < 30$) doit être interprétée dans le contexte de la distribution sous-jacente de l'IMC dans la population.
3. L'IMC est jugé fiable pour suivre les tendances pondérales à l'échelle des populations en raison de sa corrélation élevée avec le pourcentage d'adiposité. La prévalence de l'obésité définie selon l'IMC représente même une sous-estimation prudente de la population à risque.

Les conclusions du premier chapitre du rapport Fraser sont soit réfutées, soit considérablement atténuées lorsqu'on tient compte de ces problèmes. L'allégation selon laquelle « la situation canadienne ne présente pas en général une tendance déconcertante ou négative » ne nous paraît pas fondée. Nous espérons que le présent commentaire pourra servir de guide aux professionnels de la santé publique pour interpréter des analyses des tendances temporelles de l'IMC ou pour produire des analyses similaires.

MOTS CLÉS : indice de masse corporelle; poids; Canada; obésité; embonpoint; tendances

The translation of the Abstract appears at the end of this article.

Rev can santé publique 2014;105(6):e456-e459.

Un récent rapport publié par l'Institut Fraser, *Obesity in Canada: Overstated Problems, Misguided Policy Solutions*, fait de nombreuses allégations controversées qui ont de nombreuses implications au sujet de l'état de l'obésité au Canada¹. Dans le premier chapitre du rapport, on analyse les données d'enquêtes nationales de Statistique Canada pour conclure sans équivoque à l'absence d'une « tendance négative ou déconcertante » de l'obésité et de l'embonpoint. Cette conclusion étonne et intrigue à la fois, car elle s'oppose à la documentation scientifique sur les tendances pondérales canadiennes²⁻⁵.

Un examen plus attentif révèle cependant un nombre effarant de fautes graves à la fois dans les méthodes d'analyse et dans le raisonnement scientifique du rapport. Ce genre d'analyse quantitative mais bancal est trompeuse pour les lecteurs qui ne sont pas familiers avec les méthodes statistiques. Ce commentaire décrit donc brièvement trois grands problèmes de l'analyse que fait le rapport Fraser des tendances temporelles de prévalence de l'IMC au Canada. On espère qu'il servira de guide pour éviter certains écueils de l'analyse des tendances temporelles de l'IMC et qu'il aidera les professionnels de la santé à mieux évaluer les analyses semblables à celle du rapport.

Retour sur la tendance de la prévalence de l'embonpoint et de l'obésité

La figure 1 du rapport Fraser illustre la prévalence de l'embonpoint et de l'obésité chez les adultes canadiens estimée à

partir de données autodéclarées de l'IMC dans les cycles 2003 à 2012 de l'Enquête sur la santé dans les collectivités canadiennes (ESCC). Ces données sont représentées par des points noirs dans la figure 1 du présent commentaire. Le chevauchement mutuel des intervalles de confiance des estimations de la prévalence de 2008 à 2012 pousse les auteurs du rapport à conclure qu'il n'y a « aucun écart significatif entre les taux de 2008 et de 2012 », ce qui vient vraisemblablement étayer la conclusion générale selon laquelle les tendances pondérales sont « stables ou en cours de stabilisation » et qu'il n'existe « aucune tendance négative ou déconcertante » (traductions libres)¹.

Cette analyse en apparence inoffensive est parsemée d'erreurs graves. Premièrement, le chevauchement des intervalles de confiance est une méthode de comparaison statistique incorrecte qui peut mener à de fausses conclusions⁶. Par exemple, l'application d'un test z pour deux échantillons montre un écart statistiquement significatif ($p \approx 0,01$) entre les estimations de la

Affiliation de l'auteur

Institut national de santé publique du Québec; Département d'épidémiologie, de biostatistique et de santé au travail, Université McGill, Montréal (Québec)

Correspondance : Ernest Lo, Institut national de santé publique du Québec, 190, boulevard Crémazie Est, Montréal (Québec) H2P 1E2, courriel : ernest.lo@inspq.qc.ca

Remerciements : L'auteur remercie Denis Hamel, Yun Jen, Patricia Lamontagne, Sylvie Martel et Colin Steensma pour leurs commentaires sur les versions antérieures de cet article.

Conflit d'intérêts : aucun à déclarer.

Note de la rédaction : À la demande et aux frais de l'auteur, cet article est publié à la fois en anglais et en français.

Figure 1. Prévalence de l'embonpoint et de l'obésité (IMC ≥ 25), pour les deux sexes

prévalence en 2008 et 2012 et ce, malgré les intervalles de confiance qui se chevauchent.

Deuxièmement, en repérant les zones qui ont des prévalences statistiquement similaires, les auteurs font en fait des comparaisons multiples entre toutes les combinaisons de paires d'estimations. Alors que le niveau de confiance (α) d'une seule comparaison par paire (correctement effectuée) est de 95 %, le véritable seuil des tests de comparaison multiples dépend du nombre de points de données dans la série chronologique, de leurs variances, ainsi que du nombre de points dans la zone d'intérêt et de sa probabilité de sélection ultérieure. Le fait de ne trouver aucun écart dans les comparaisons multiples sur la période 2008–2012 est donc associé à un niveau de signification statistique inconnu.

Troisièmement, la fréquence des différences significatives observées pour les comparaisons deux à deux est sensible aux changements de la taille de l'échantillon au fil du temps. Comme on peut le voir à la figure 1, les intervalles de confiance pour les dernières années (2007–2012) sont 60 % plus larges en moyenne. Ceci s'explique par les plus petites tailles d'échantillons ($\approx 55\,000$ adultes) des estimations à mi-cycle de l'ESCC utilisées au cours de cette période comparativement aux tailles d'échantillons ($\approx 116\,000$ adultes) des cycles complets de 2003 et 2005. En conséquence, il y a une augmentation de la probabilité d'obtenir des estimations statistiquement similaires durant cette période plus récente, qui s'explique par un changement temporel de la taille d'échantillon plutôt que par un phénomène de stabilisation. Pour le démontrer, on a ajouté sur la figure 1 des points en gris qui représentent les estimations de prévalence à partir des cycles complets de l'ESCC en 2007–2008, 2009–2010 et 2011–2012. Les comparaisons de deux échantillons par le test z montrent qu'une fois que l'on a standardisé les tailles d'échantillons de toutes les estimations, les prévalences de l'embonpoint et de l'obésité ont augmenté de façon significative tous les 4 ans sur l'ensemble de la période de 2003 à 2012.

Quatrièmement, l'espacement des intervalles de deux ans entre 2003 et 2007 pour tous les tracés des tendances temporelles de prévalence du rapport Fraser, y compris la figure 1, a été rendu égal à celui des intervalles d'un an entre 2007 et 2012, ce qui

revient à comprimer l'axe des x entre les années 2003 et 2007 et à créer l'illusion visuelle que les courbes s'aplanissent les dernières années par rapport aux années antérieures. L'axe des x dans la figure 1 du présent commentaire représente correctement l'espacement de deux ans entre les cycles 2003, 2005 et 2007 de l'ESCC, ce qui donne une représentation non biaisée des tendances.

Les comparaisons multiples ne sont pas en général une méthode fiable pour évaluer les tendances de séries chronologiques, puisque de longues séquences de points de données peuvent être statistiquement similaires en raison de la variabilité des mesures individuelles et ce, même en présence d'une forte tendance⁷. Les méthodes de régression, par contre, permettent une évaluation statistiquement rigoureuse des tendances pour tous les points de données collectivement, tout en tenant compte de l'évolution de leur variance en raison de la taille de l'échantillon. Une simple régression linéaire ajustée aux données originales, à l'aide de pondérations pour tenir compte de l'effet des différentes tailles d'échantillons des cycles de l'ESCC, donne une pente positive dont la valeur estimée de 0,37 est significative ($p < 2 \times 10^{-6}$, $R^2 \approx 0,98$), ce qui indique une tendance linéaire en croissance constante dans les données. La ligne ajustée (en pointillés dans la figure 1 du présent commentaire) montre que l'ajustement est bon. Cette tendance à la hausse demeure quand on ajuste seulement les données récentes (2008–2012), période pour laquelle l'analyse de régression donne une pente d'une valeur statistiquement similaire de 0,34 ($p < 0,03$, $R^2 \approx 0,8$).

La stabilité temporelle de la prévalence de l'embonpoint cache un changement continu dans la distribution de l'IMC

La prévalence de la catégorie de l'embonpoint de l'IMC chez les Canadiens adultes est restée à peu près constante de 2003 à 2012⁸. Ces tendances sont présentées, pour les deux sexes combinés et séparés, dans les figures 2, 5 et 8 du rapport Fraser comme des arguments importants en faveur de la stabilité temporelle du poids dans la population canadienne. Cette interprétation des tendances de prévalence de l'embonpoint est à la fois naïve et trompeuse, car la stabilité de la prévalence de

Figure 2. Tendances temporelles de l'IMC moyen et par percentile pour les adultes canadiens des deux sexes combinés

l'embonpoint est en fait la conséquence d'un changement temporel soutenu et inquiétant dans la distribution de l'IMC dans la population générale.

Pour le démontrer, nous avons extrait la moyenne et les percentiles de la distribution de l'IMC pour les adultes canadiens (des deux sexes combinés) de ces mêmes cycles de l'ESCC. On peut voir aux figures 2a et 2b que les IMC moyen et médian (au 50^e percentile) de la population augmentent depuis 2003, ce qui indique un déplacement soutenu de la distribution de l'IMC vers la droite (vers des valeurs plus élevées). Ce déplacement de la distribution s'étend sur l'ensemble des valeurs de l'IMC définissant l'embonpoint, comme on le voit aux figures 2c à 2e avec les 60^e, 70^e et 80^e percentiles tracés dans le temps.

La raison pour laquelle la catégorie de prévalence de l'embonpoint est paradoxalement restée stable durant toute cette période est que la distribution de l'IMC s'élargit en plus de se déplacer. Ce phénomène est illustré à la figure 3, qui montre les distributions lissées de l'IMC estimées à partir des percentiles, pour 2003 (en noir) et 2012 (en gris). L'IMC est représenté sur l'axe des x, et les lignes pointillées verticales délimitent la

Figure 3. Distributions estimées de l'IMC pour 2003 et 2012

catégorie de l'embonpoint; les aires sous les courbes de distribution comprises dans cet intervalle sont égales à la proportion ou à la prévalence de personnes en embonpoint dans les populations respectives. On peut voir que l'augmentation de la part des personnes en embonpoint dans la population en raison du déplacement vers la droite de la distribution « A » est compensée par l'élargissement de la distribution « B » causé par l'augmentation proportionnelle des personnes obèses. La stabilité temporelle de la prévalence de la catégorie d'IMC dénotant un embonpoint représente ainsi la redistribution continue des personnes qui s'éloignent de la catégorie du poids normal de la distribution de l'IMC pour entrer dans celle de l'obésité.

Évaluation de l'impact des limites de l'IMC sur les tendances temporelles

Le rapport Fraser explique ensuite l'incapacité de l'IMC à distinguer la masse maigre de la masse adipeuse. Les auteurs laissent alors sous-entendre que cette importante limite réfute les analyses habituelles des tendances temporelles de l'IMC tout en étayant les conclusions du premier chapitre. Tous les indicateurs de santé présentent cependant des limites; cela ne veut pas nécessairement dire qu'un indicateur donné ne convient pas à l'application à laquelle il est destiné. Il s'avère qu'en raison de sa corrélation élevée avec les mesures directes du pourcentage d'adiposité⁹, l'IMC est recommandé pour surveiller les tendances temporelles de la prévalence du risque dû à une adiposité élevée à l'échelle d'une population⁹⁻¹².

L'obésité définie selon l'IMC est un indicateur de faible sensibilité, mais de forte spécificité^{9,11} pour l'adiposité élevée, ce qui entraîne une sous-estimation de la population à risque. Ce phénomène est aggravé par l'important biais systématique vers le bas de l'IMC autodéclaré, estimé à $\approx 7-8\%$ dans l'ESCC de 2008 pour les adultes de 18 à 74 ans¹³, mais dont l'ampleur dépend en général de l'enquête¹⁴. Par ailleurs, le déplacement de l'adiposité vers le haut observé dans la catégorie de l'IMC dénotant l'obésité¹⁵ indique que les estimations du taux d'augmentation temporelle de la population à risque pourraient aussi être sous-estimées. En raison de ces principales limites de l'IMC, la progression observée de l'obésité d'après les données de l'ESCC est ainsi une mesure conservatrice des niveaux de risque pour la santé dus au gain de poids dans la population canadienne, ce qui

tend à contredire encore davantage les conclusions du chapitre du rapport Fraser.

CONCLUSIONS

Nous avons présenté trois problèmes dans l'analyse des tendances temporelles de l'IMC tirée du premier chapitre du rapport Fraser¹ :

1. Repérer les chevauchements des intervalles de confiance est une méthode statistiquement imparfaite d'évaluation des tendances; les méthodes de régression, dont on peut faire l'interprétation statistique sans ambiguïté sont préférables. En effet, elles tiennent vraiment compte des changements temporels dans la taille de l'échantillon et mesurent le comportement des données dans leur ensemble.
2. La stabilité temporelle dans la prévalence de l'embonpoint ($25 \leq \text{IMC} < 30$) est en fait la conséquence d'un changement soutenu dans la distribution générale de l'IMC dans la population.
3. L'IMC est jugé fiable pour suivre les tendances pondérales à l'échelle des populations en raison de sa corrélation élevée avec le pourcentage d'adiposité. La prévalence de l'obésité, estimée à partir de l'IMC, est une sous-estimation conservatrice de la population à risque en raison d'une adiposité élevée.

Les résultats et l'interprétation d'une analyse des tendances temporelles de l'IMC peuvent être très différents si ces problèmes ne sont pas pris en compte. En particulier, de nombreuses constatations du premier chapitre du rapport Fraser en sont soit réfutées, soit considérablement atténuées. Il semble que le taux d'augmentation de la prévalence de l'obésité est inférieur à ce qu'il a été par le passé, et qu'il y a eu une curieuse fluctuation vers le bas de la prévalence estimée en 2012 chez les hommes au Canada. Toutefois, les données existantes, lorsqu'elles sont soumises à une analyse plus rigoureuse, ne présentent pas l'absence sans équivoque d'une « tendance déconcertante ou négative » alléguée par les auteurs du rapport. De façon générale, nous espérons que le présent commentaire contribuera à guider les professionnels de la santé publique dans leur interprétation des analyses des tendances temporelles de l'IMC ou dans la réalisation d'analyses similaires.

RÉFÉRENCES

1. Esmail N, Basham P, *Obesity in Canada: Overstated Problems, Misguided Policy Solutions*, Institut Fraser, 2014. (<http://www.fraserinstitute.org>)
2. *Obésité au Canada*, Ottawa, Agence de la santé publique du Canada et Institut canadien d'information sur la santé, 2011.
3. Gotay CC, Katzmarzyk PT, Janssen I, Dawson MY, Aminoltejeri K, Bartley NL, 2012, « Updating the Canadian obesity maps: An epidemic in progress », *Rev can santé publique*, 104, 1, p. e64-e68.
4. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C et coll., 2014, « Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: A systematic analysis for the Global Burden of Disease Study 2013 », *Lancet*, 384, 9945, p. 766-781.
5. Twells L, Gregory D, Reddigan J, Midodzi WK, 2014, « Current and predicted prevalence of obesity in Canada: A trend analysis », *JAMC Ouvert*, 2, 1, p. E18.

6. Schenker N, Gentleman JF, 2001, « On judging the significance of differences by examining the overlap between confidence intervals », *Am Stat*, 55, 3, p. 182-186.
7. Brockwell PJ, Davis RA, *Introduction to Time Series and Forecasting*, 2^e édition, New York, Springer, 2002.
8. Statistique Canada, « Embonpoint et obésité chez les adultes (mesures autodéclarées), 2011 ». Sur Internet : <http://www.statcan.gc.ca/pub/82-625-x/2012001/article/11664-fra.htm> (Consulté en version anglaise le 1^{er} novembre 2014.)
9. Romero-Corral A, Somers VK, Sierra-Johnson J, Thomas RJ, Collazo-Clavell ML, Korinek J et coll., 2008, « Accuracy of body mass index in diagnosing obesity in the adult general population », *Int J Obes (Lond)*, 32, 6, p. 959-966.
10. Centers for Disease Control and Prevention, *Body Mass Index: Considerations for Practitioners*, 2014. Sur Internet : <http://www.cdc.gov/obesity/downloads/bmiforpractitioners.pdf> (Consulté le 1^{er} novembre 2014.)
11. Javed A, Jumean M, Murad MH, Okorodudu D, Kumar S, Somers VK et coll., 2014, « Diagnostic performance of body mass index to identify obesity as defined by body adiposity in children and adolescents: A systematic review and meta-analysis », *Pediatr Obes*, 10. L'article a d'abord été publié en ligne : 24 JUN 2014, DOI : 10.1111/ijpo.242.
12. Townsend N, *Obesity and Overweight Surveillance in England: What is measured and where are the gaps?*, Oxford (Royaume-Uni), National Obesity Observatory, 2009.
13. Shields M, Connor GS, Janssen I, Tremblay MS, 2011, « Biais dans les estimations autodéclarées de l'obésité dans les enquêtes canadiennes sur la santé : le point sur les équations de correction applicables aux adultes », *Rapports sur la santé*, 22, 3, p. 35-45.
14. Connor GS, Tremblay MS, 2010, « The bias in self-reported obesity from 1976 to 2005: A Canada-US comparison », *Obesity* (Silver Spring), 18, 2, p. 354-361.
15. Shields M, Tremblay MS, Connor GS, Janssen I, 2012, « Mesures de l'obésité abdominale à l'intérieur des catégories d'indice de masse corporelle, 1981 et 2007-2009 », *Rapports sur la santé*, 23, 2, p. 33-38.

Reçu : 30 mai 2014

Accepté : 13 octobre 2014

ABSTRACT

The first chapter of the Fraser report "Obesity in Canada: Overstated Problems, Misguided Policy Solutions" presents a flawed and misleading analysis of BMI time trends. The objective of this commentary is to provide a tutorial on BMI time trend analysis through the examination of these flaws.

Three issues are discussed:

1. Spotting regions of confidence interval overlap is a statistically flawed method of assessing trend; regression methods which measure the behaviour of the data as a whole are preferred.
2. Temporal stability in overweight ($25 \leq \text{BMI} < 30$) prevalence must be interpreted in the context of the underlying population BMI distribution.
3. BMI is considered reliable for tracking population-level weight trends due to its high correlation with body fat percentage. BMI-defined obesity prevalence represents a conservative underestimate of the population at risk.

The findings of the Fraser report Chapter 1 are either refuted or substantially mitigated once the above issues are accounted for, and we do not find that the 'Canadian situation largely lacks a disconcerting or negative trend', as claimed. It is hoped that this commentary will help guide public health professionals who need to interpret, or wish to perform their own, time trend analyses of BMI.

KEY WORDS: Body mass index; body weight; Canada; obesity; overweight; trends