

## Enquête sur la santé dans les collectivités canadiennes, 2009-2010

### Note méthodologique sur la pertinence de produire des estimations régionales à partir des données de la première année du cycle 2009-2010

Institut de la statistique du Québec  
Direction de la méthodologie et de la qualité  
Nathalie Plante et Robert Courtemanche  
Novembre 2010

La pertinence de produire des estimations à partir des données de mi-collecte de l'ESCC relève de deux aspects méthodologiques : l'exactitude et la précision. La précision au sens statistique est une mesure de l'écart entre une estimation obtenue à partir de l'échantillon et le résultat qu'on aurait obtenu d'un recensement, alors que l'exactitude désigne l'écart entre un résultat de l'enquête et la valeur réelle (mais inconnue) d'une caractéristique de la population.

#### **Les estimations régionales produites à partir des données de la première année d'un cycle sont en général exactes, mais imprécises.**

Elles sont exactes puisque :

- l'échantillon de la première année de collecte est représentatif au niveau de la région sociosanitaire, c'est-à-dire que les statistiques pondérées issues de cet échantillon peuvent être inférées à la population de chaque région ou à l'ensemble de la population visée ;
- la pondération permet de tenir compte des probabilités de sélection variables et de minimiser les biais potentiels dus, entre autres, à la non-réponse.

Elles sont en général imprécises puisque :

- le nombre de répondants par région sociosanitaire est faible ; il en résulte une précision des estimations et une puissance statistique limitées ;
- les objectifs de précision sur lesquels reposent les tailles d'échantillon fixées sont basés sur un cycle complet, soit deux années de collecte ; après une année de collecte seulement, on ne peut prétendre atteindre les objectifs de précision visés pour le cycle 2009-2010 ;
- une analyse portant sur les répondants de l'année 2009 seulement permet, grosso modo, d'obtenir une bonne précision (soit un coefficient de variation de 15 % ou moins) pour une proportion de l'ordre de 15 % ou plus portant sur l'ensemble de la population de la plupart des régions ; pour une proportion inférieure à 15 %, ou pour l'étude d'une sous-population d'une région, la précision est en général plus faible. (Note : Cet ordre de grandeur est fourni à titre indicatif seulement. L'utilisateur doit se référer aux tableaux spécifiques à chaque région fournis par Statistique Canada pour une évaluation plus juste. Certaines régions où le nombre de répondants est sensiblement plus important peuvent obtenir une meilleure précision.)

Dans ce contexte, on peut se questionner sur la pertinence de produire des statistiques régionales à mi-collecte sachant que :

- les écarts qui prévalent dans la population entre deux années consécutives sont faibles en général, de sorte qu'étant donné la précision attendue, l'estimation pour l'année 2009 risque fort d'apporter peu d'information supplémentaire par rapport au cycle 2007-2008 ;
- les comparaisons qui peuvent être faites au niveau régional ou temporel devront impérativement tenir compte de la précision des estimations et sont sujettes à une faible puissance statistique ;
- les efforts investis risquent de donner lieu à peu de résultats significatifs.

À partir des données des cycles 2.1 (2003) et 2007-2008 de l'ESCC, les tableaux 1 et 2 illustrent la problématique de l'interprétation des statistiques issues des données de mi-collecte. Le tableau 1 montre que l'analyse des données de mi-collecte de l'année 2007 ou de l'année 2008 ne permet pas de détecter d'écart significatif entre la région de l'Outaouais et l'ensemble du Québec pour les deux indicateurs étudiés, alors que l'analyse des données du cycle complet permet de démontrer que l'Outaouais diffère de l'ensemble du Québec pour ces mêmes indicateurs.

Tableau 1. Participation et limitation des activités et statut de fumeurs - comparaison de la région de l'Outaouais et de la province de Québec pour les années 2007, 2008 et 2007-2008 (combinées) de l'ESCC

<b>Indicateurs</b>	<b>2007<sup>a</sup></b>		<b>2008<sup>a</sup></b>		<b>2007-2008<sup>b</sup></b>	
	Outaouais %	Québec %	Outaouais %	Québec %	Outaouais %	Québec %
Population de 12 ans et plus						
Participation et limitation des activités	28,53	24,46	25,00	22,75	26,75	23,60
Fumeurs réguliers	23,19	19,75	22,24	18,48	22,71	19,12

Source : Fichier de partage de l'ESCC, cycle 2007-2008.

a. Absence de différence significative au seuil 0,05 entre l'Outaouais et le reste du Québec.

b. Différence significative au seuil 0,05 entre l'Outaouais et le reste du Québec.

Le tableau 2 illustre quant à lui la précision sensiblement plus faible observée pour la prévalence de la population de 12 ans et plus de la région de la Capitale-Nationale atteinte d'un trouble d'anxiété lorsque la statistique repose sur un échantillon de mi-collecte. Ainsi, pour cette région, il n'est pas possible de déclarer significatifs (au seuil 0,05) des écarts observés pourtant non négligeables entre le cycle 2.1 (2003) et l'année 2007 ou 2008 (données de mi-collecte – écarts observés non significatifs respectifs de 1,52 % et 1,56 % avec le cycle 2.1), alors que l'on peut démontrer un écart significatif avec le cycle 2007-2008 (cycle complet – écart observé significatif de 1,54 % avec le cycle 2.1).

Tableau 2. Prévalence de la population de 12 ans et plus de la région de la Capitale-Nationale atteinte d'un trouble d'anxiété - tests de différence entre les années 2003 et 2007-2008 (2007 seulement, 2008 seulement ou les deux années combinées) de l'ESCC.

<b>Trouble d'anxiété Population de 12 ans et plus Région de la Capitale-Nationale</b>	<b>2003</b>	<b>2007</b>	<b>2008</b>	<b>2007-2008</b>
Proportion (%)	3,62	5,14	5,18	5,16
Erreur-type	0,40	0,84	0,90	0,61
Coefficient de variation (%)	11,0	16,3	17,4	11,8
Seuil observé du test de différence avec le cycle 2.1 (2003)		0,08	0,08	0,03

Source : Fichiers de partage de l'ESCC, cycles 2.1 (2003) et 2007-2008.

Mises à part les préoccupations de précision et de puissance statistiques, il importe de s'assurer de la comparabilité des indicateurs étudiés avant de conclure à des écarts significatifs. Bien qu'à ce jour, une analyse complète n'ait pas été effectuée pour le cycle 2009-2010, l'examen des nombres de répondants de ce cycle porte à croire que la comparabilité régionale intracycle ne posera pas problème. Par comparaison régionale intracycle, on entend la comparaison des régions entre elles ou d'une région avec l'ensemble des autres régions pour le cycle 2009-2010, ou encore pour une année de ce cycle seulement. À l'opposé, s'il doit y avoir des comparaisons temporelles des données de mi-collecte du cycle 2009-2010 avec les données des cycles antérieurs en dépit de la faible puissance statistique offerte par une telle analyse, la question de la comparabilité se pose et des précautions analytiques s'imposent.