

Pondérations de l'échantillonnage, pondérations de l'estimation et ajustements de non-réponse

Lorsqu'un échantillon est choisi pour un sondage avec des probabilités inégales (comme ce fut le cas pour l'échantillon du questionnaire détaillé du SNM), des pondérations sont utilisées pour faire des estimations afin que l'échantillon pondéré soit représentatif de la population. Les pondérations de l'échantillonnage original (sur la base des probabilités de sélection) sont habituellement ajustées avant l'estimation afin de réduire les biais possibles de non-réponse. Les recensements (un recensement fut entrepris pour les questions de base du SNM 2007) sont sujets aux absences de réponse et, conséquemment, les pondérations à utiliser pour l'estimation peuvent être calculées afin de réduire les biais possibles de non-réponse. Dans le SNM, les ajustements de non-réponse pour la composante échantillon du sondage et la composante recensement ont été faits au niveau provincial par type de médecins, par groupe d'âge et par sexe à l'aide de la méthode de calibration (référence : Méthodes et pratiques d'enquête. Statistique Canada catalogue no.12-587-XPF, 2003.) Le Fichier maître du SNM a servi de population de référence pour cette calibration.

Population admissible : Des 62 441 médecins apparaissant sur la liste initiale, l'admissibilité a pu être déterminée pour 20 869 médecins, dont 1 630 ont été jugés non admissibles. Le processus de pondération et d'ajustement pour la non-réponse comprenait les 19 239 médecins qui ont répondu au sondage et les 1 630 jugés non admissibles, ce qui a permis l'estimation du nombre de médecins non admissibles parmi les 41 572 médecins dont l'admissibilité n'a pas été confirmée. Cette méthode produit une estimation de 57 757 médecins admissibles. Les cellules ombragées dans le tableau ci-dessous sont les nombres estimés.

Estimation de la population admissible, SNM 2007 :

	Admissibilité déterminée	Admissibilité non déterminée	Total
Admissibles	19 239	38 518	57 757
Non admissibles	1 630	3 054	4 684
Total	20 869	41 572	62 441

Échantillon de répondants : Pour les questions de base du sondage, nous avons reçu 19 239 réponses représentant les 57 757 médecins admissibles. Après les ajustements pour les absences de réponse dans les différents groupes démographiques, les pondérations définitives pour ces réponses se situent en moyenne assez exactement à 3 (3,002), variant entre 1,00 et 4,31.

Pour les questions à partir des questionnaires détaillés, nous avons reçu 7 347 réponses représentant 57 757 médecins admissibles. Après les ajustements pour les absences de réponse dans les différents groupes démographiques, les pondérations définitives pour ces réponses se situent en moyenne à 7,86, variant entre 1,00 et 13,72.

Pour plus de détails sur la répartition des pondérations, consultez L'ANNEXE C.

Utilisation des pondérations

Il y a deux ensembles de pondérations : un pour les questions apparaissant sur le questionnaire de base et le questionnaire détaillé (questions de base) et un pour les questions apparaissant seulement sur les questionnaires détaillés (questions détaillées). Pour les estimations impliquant seulement les questions de base, comme 'Milieu principal de soins aux patients par province', ou 'Satisfaction professionnelle selon le Milieu principal de soins aux patients', la pondération devrait être utilisée pour la question de base. De façon similaire, les estimations qui impliquent seulement les questions détaillées, comme 'Le type d'accès à l'Internet pour le système de tenue de dossiers', la pondération devrait être utilisée pour la question détaillée. Pour les tableaux qui utilisent les questions de base et les questions détaillées, comme 'Le type d'accès à l'Internet selon le milieu principal de soins aux patients', la pondération devrait être utilisée pour la question détaillée.

Les deux ensembles de pondération permettent de produire deux estimations pour l'une ou l'autre des questions de base. L'une est l'estimation naturelle, utilisant tous les répondants et les pondérations pour les questions de base, et l'autre est obtenue en utilisant les réponses aux questions de base des répondants au questionnaire détaillé (les questions de base forment un sous-ensemble des questions apparaissant sur le questionnaire plus long et détaillé) et les pondérations pour les questions détaillées. (Les chiffres marginaux pour la question de base à partir d'un tableau d'une question de base par une question détaillée seront des estimations du second type.) Ces deux estimations seront légèrement différentes parce qu'elles reposent sur deux ensembles différents de répondants. Les estimations basées sur l'échantillon plus large (échantillon du questionnaire de base) seront plus précises à cause du plus grand nombre de répondants.

Variabilité d'échantillonnage des estimations

Les données provenant du SNM 2007 sont basées sur un sondage par échantillonnage dans le cas des questions détaillées et sur un recensement avec un nombre considérable d'absences de réponse dans le cas des questions de base. Des résultats différents auraient été obtenus si un recensement avec absence de non-réponse avait été mené pour toutes les questions. Ces différences sont appelées erreurs d'échantillonnage et erreurs de non-réponse, et la théorie d'échantillonnage nous donne un moyen d'estimer leur taille. Pour le SNM, il a été présumé que la non-réponse était approximativement aléatoire, alors elle peut être traitée essentiellement comme partie du processus de sélection au hasard.

La variabilité de l'estimation d'un sondage dépend de nombreux facteurs, notamment la taille de l'échantillon, sa répartition entre les strates, la taille de l'estimation et les détails du plan d'échantillonnage. À l'aide de la formule usuelle pour la variance d'échantillonnage des estimations à partir d'échantillons aléatoires simples [réf. : Cochran, W. (1977). *Sampling Techniques*. John Wiley and Sons, New York] et un ajustement conservateur de l'effet du plan de sondage pour expliquer la stratification et la calibration, des principes de variabilité peuvent être établis pour le SNM 2007 et

certains énoncés généraux peuvent être formulés concernant la précision des estimations :

- Les estimations utilisant les questions de base du SNM pour des proportions de l'ensemble de la population de médecins dans toutes les provinces se situeront à l'intérieur de 1 point de pourcentage de la proportion véritable, 19 fois sur 20.
- Les estimations à l'échelle du Canada pour les questions de base, réservées aux sous-populations d'Autres spécialistes ou de Médecins de famille/omnipraticiens, se situeront à l'intérieur de 1,5 point de pourcentage, 19 fois sur 20.
- Les estimations à l'échelle du Canada pour les questions détaillées pour l'ensemble de la population de médecins se situeront à l'intérieur de 3,2 points de pourcentage, 19 fois sur 20.
- Les estimations à l'échelle du Canada pour les questions détaillées posées aux sous-populations d'Autres spécialistes ou de Médecins de famille/omnipraticiens se situeront à l'intérieur 5 points de pourcentage 19 fois sur 20.

Ce sont des principes conservateurs, basés sur la taille de la population et de l'échantillon, et sur la proportion la plus difficile à estimer. Cette proportion est de 50 %; pour les estimations plus faibles ou plus grandes, les intervalles de confiance sont considérablement rétrécis. Par exemple, on peut s'attendre qu'une estimation de 5 % (ou de 95 %) pour tous les médecins de l'ensemble du Canada se situera à l'intérieur de 0,45 points de pourcentage de la valeur réelle, 19 fois sur 20.

L'amplitude de ces intervalles de confiance est très sujette aux tailles de la population et de l'échantillon. Les amplitudes de l'intervalle de confiance pour les estimations de Terre-Neuve-et-Labrador ou de l'île-du-Prince-Édouard sont environ dix fois celles du Canada, alors que celles pour le Québec ou l'Ontario sont environ trois fois seulement celles pour le Canada. Veuillez vous référer à l'ANNEXE A pour des détails supplémentaires concernant les limites de confiance et l'ANNEXE B pour une liste des limites de confiance provinciales/territoriales.

La validité des estimations de la variabilité de l'échantillonnage, ainsi que les intervalles de confiance et les tests d'hypothèse qui en résultent sont fonction de la validité des hypothèses sur lesquelles elles sont basées. L'hypothèse essentielle, c'est que l'effet combiné de l'échantillonnage et de la non-réponse est à peu près semblable à un échantillon aléatoire à l'intérieur des classes qui ont été utilisées pour la calibration. C'est également l'hypothèse selon laquelle les estimations du SNM 2007 vont refléter avec exactitude l'entière population des médecins admissibles.

Nous disposons de peu d'éléments de données pour mesurer cette hypothèse mais, pour les caractéristiques des médecins qui étaient connues tant chez les répondants que les non-répondants, les deux groupes paraissent assez semblables (voir section précédente). La prochaine section présente les résultats d'un sondage supplémentaire réalisé auprès des non-répondants au questionnaire détaillé du SNM 2007, lequel a décelé des différences minimales mais statistiquement significatives entre les médecins répondants et non répondants.

La pondération corrige pour la surreprésentation ou la sous-représentation des groupes définis par province, par type de médecins (spécialiste vs MF/omni), par âge et par sexe. Par conséquent, tout biais dans les réponses, attribuable à la non-réponse différentielle entre ces groupes, a été supprimé par la calibration de ces pondérations.