



Numéro 92-394-XIF au catalogue

Couverture

Rapport technique du recensement de 2001



Statistique
Canada

Statistics
Canada

Canada

PUBLICATIONS ÉLECTRONIQUES DISPONIBLES À
www.statcan.ca



Couverture

	Page
INTRODUCTION.....	4
1. COLLECTE DES DONNÉES DU RECENSEMENT	5
1.1 Généralités.....	5
1.2 Méthodes de collecte.....	5
2. TRAITEMENT DES DONNÉES DU RECENSEMENT	7
2.1 Généralités.....	7
2.2 Dépouillement régional	8
2.3 Imagerie	9
2.4 Vérification interactive.....	9
2.5 Codage automatisé et codage interactif.....	10
2.6 Contrôle et imputation.....	11
2.7 Ajustements de la couverture en fonction des logements inoccupés et des logements non répondants.....	11
2.8 Pondération.....	11
3. ÉVALUATION DE LA QUALITÉ DES DONNÉES	13
3.1 Généralités.....	13
3.2 Causes des erreurs de couverture et mesures prises en vue d'en réduire le nombre.....	13
3.3 Erreurs de couverture – Définitions	14
4. UNIVERS DU RECENSEMENT	16
4.1 Généralités.....	16
4.2 Univers de la population	16
4.3 Univers des logements	17
4.4 Univers des ménages	18
4.5 Univers des familles de recensement.....	18
4.6 Univers des familles économiques	18
4.7 Liens existant entre les univers	19
5. SOUS-DÉNOMBREMENT NET AU RECENSEMENT DE 2001	20
5.1 Sous-dénombrement net de la population.....	24
6. PROGRAMME DE MESURE DE L'ERREUR DE COUVERTURE	26
6.1 Portée et objectifs	26
6.2 Améliorations	27
7. ÉTUDE SUR LA CLASSIFICATION DES LOGEMENTS.....	29
7.1 Introduction	29
7.2 Méthodologie	29
7.2.1 Stratification et sélection de l'échantillon	29
7.2.2 Interviews sur le terrain.....	31
7.2.3 Dépouillement, codage et contrôle	31

	Page
7.2.4 Non-réponse, imputation et pondération	32
7.3 Résultats	33
7.3.1 Logements inscrits comme étant inoccupés au recensement.....	33
7.3.2 Logements inscrits comme étant non répondants au recensement de 2001.....	39
8. CONTRE-VÉRIFICATION DES DOSSIERS	41
8.1 Introduction	41
8.2 Méthodologie – Préparation de l'échantillon.....	41
8.3 Méthodologie – Collecte des données	43
8.3.1 Cadre de la collecte des données	43
8.3.2 Opérations de collecte	44
8.3.3 Dépistage.....	46
8.3.4 Statistiques sur la collecte	47
8.3.5 Enquête de suivi de la Contre-vérification des dossiers.....	48
8.4 Méthodologie – Traitement des adresses et classification des personnes choisies	50
8.5 Méthodologie – Estimation	54
8.6 Méthodologie – Production des tableaux	57
9. ÉTUDE PAR APPARIEMENT AUTOMATISÉ	59
9.1 Méthodologie	59
9.2 Plan d'échantillonnage et traitement.....	59
10. ÉTUDE SUR LES LOGEMENTS COLLECTIFS.....	62
10.1 Méthodologie	62
10.2 Traitement.....	62
11. ESTIMATIONS DE L'ERREUR DE COUVERTURE.....	64
11.1 Méthodologie	64
11.2 Résultats	65
11.2.1 Sous-dénombrement brut.....	65
11.2.2 Surdénombrement	65
11.2.3 Sous-dénombrement net	66
12. COMPARABILITÉ CHRONOLOGIQUE	71
12.1 Comparaisons avec les chiffres publiés du recensement de 2001	71
13. ÉVALUATION DE LA CONTRE-VÉRIFICATION DES DOSSIERS.....	77
13.1 Introduction	77
13.2 Comparaisons avec les chiffres publiés du recensement de 2001	77
13.2.1 Personnes dénombrées.....	77
13.3 Comparaison avec les estimations de la population	79
13.3.1 Personnes décédées	79
13.3.2 Migration interprovinciale.....	80
13.4 Composantes de la croissance démographique	80
14. RÉSERVES ET ÉTABLISSEMENTS INDIENS AYANT OPPOSÉ UN REFUS.....	82
14.1 Introduction	82
14.2 Méthodologie	82
14.3 Résultats en 2001	83

	Page
ANNEXES	
Annexe A. Sigles et acronymes	85
Annexe B. Glossaire	86
Annexe C. Produits et services du recensement de 2001	87
BIBLIOGRAPHIE.....	88

Introduction

Le recensement de 2001 a nécessité la participation de l'ensemble de la population canadienne, soit environ 30 millions de personnes réparties sur un territoire couvrant 9 millions de kilomètres carrés. Même si des normes de qualité rigoureuses régissent la collecte et le traitement des données, il est impossible de supprimer toutes les erreurs. Afin d'aider les utilisateurs à évaluer l'utilité des données du recensement en fonction de leurs besoins, la série des Rapports techniques du recensement de 2001 explique le cadre théorique et les définitions ayant servi à mener le recensement, de même que les méthodes de collecte des données et les méthodes de traitement employées. De plus, elle couvre les principales sources d'erreur, y compris, dans la mesure du possible, la taille de ces erreurs, ainsi que les circonstances inhabituelles pouvant limiter l'utilité ou l'interprétation des données du recensement. À l'aide de ces renseignements, les utilisateurs peuvent évaluer les risques entourant des conclusions ou des décisions fondées sur les données du recensement.

Ce Rapport technique du recensement de 2001 porte sur les erreurs de couverture, qui sont survenues lorsque des personnes ou des logements ont été omis lors du recensement de 2001 ou dénombrés par erreur. Les erreurs de couverture constituent l'un des plus importants types d'erreur lors du recensement, étant donné qu'elles touchent non seulement la précision des chiffres des divers univers du recensement mais aussi la précision de toutes les données du recensement portant sur les caractéristiques de ces univers. **Les utilisateurs des données du recensement devraient savoir que la présence d'erreur de couverture au recensement de 2001 signifie que les totalisations peuvent présenter les résultats d'un dénombrement incomplet. Cette mise en garde vaut particulièrement au sein des groupes qui présentent de fortes erreurs de couverture. Les utilisateurs sont priés de voir les chapitres 5, 7 et 11 pour obtenir des estimations de l'ampleur de l'erreur de couverture de la population et des ménages selon divers niveaux et groupes démographiques et géographiques.**

Les chapitres 1 à 3 décrivent les procédures de collecte et de traitement de données au cours desquelles les erreurs de couverture peuvent survenir, ainsi que les démarches entreprises en vue de contrôler de telles erreurs. Le chapitre 4 définit les univers que le recensement tente de couvrir et sert à encadrer le Programme de mesure de l'erreur de couverture. Le chapitre 5 présente des estimations du sous-dénombrement net au recensement de 2001. Le chapitre 6 donne une vue d'ensemble du Programme de mesure de l'erreur de couverture pour le recensement de 2001, alors que les chapitres 7 à 10 décrivent la méthodologie et les résultats de chacune des études sur la couverture. Le chapitre 11 montre de quelle façon les résultats des quatre études sont jumelés pour former les estimations de l'erreur de couverture et donne d'autres estimations de l'erreur de couverture du recensement de 2001. Le chapitre 12 fournit une perspective chronologique et le chapitre 13 présente l'évaluation de la plus vaste étude sur la couverture : la Contre-vérification des dossiers. Enfin, le chapitre 14 présente un aperçu des questions liées à la couverture dans le contexte des réserves indiennes. Veuillez prendre note que la liste de toutes les abréviations utilisées dans le présent rapport figure à l'annexe A.

Le présent rapport a été préparé par Mark Armstrong, Peter Dick, Heather Farr, Gildas Kleim, Josée Morel, Heather Richards, Karen Switzer, Alain Théberge et Christian Thibault de la Division des méthodes d'enquêtes sociales. Grâce aux précieux commentaires qu'ils ont formulés sur les versions préliminaires, Jean Dumais, David Dolson et Jocelyn Tourigny ont contribué à améliorer le contenu et la lisibilité du rapport final. Nous remercions également de leur précieux concours le personnel de la Division des opérations du recensement, de la Division de la démographie et de la Division des méthodes d'enquêtes sociales.

Les utilisateurs trouveront des renseignements supplémentaires sur les concepts, les variables et la géographie du recensement dans le *Dictionnaire du recensement de 2001* (n° 92-378-XIF au catalogue), de même qu'un aperçu des différentes étapes du recensement dans *Le recensement de 2001 en bref* (n° 92-379-XIF au catalogue).

1. Collecte des données du recensement

1.1 Généralités

L'étape de la collecte des données du processus du recensement vise à faire en sorte que les quelque 11,8 millions de ménages du Canada soient dénombrés le jour du recensement (le mardi 15 mai au recensement de 2001). Le recensement vise l'ensemble de la population du Canada, qui comprend les citoyens canadiens (par naissance ou par naturalisation), les immigrants reçus et les résidents non permanents. Les résidents non permanents sont les personnes au Canada qui sont titulaires d'un permis de travail, d'un permis de séjour pour étudiants ou d'un permis ministériel, ou qui revendiquent le statut de réfugié, ainsi que les membres de leur famille qui vivent avec eux.

Le recensement dénombre également les citoyens canadiens et les immigrants reçus qui sont temporairement à l'extérieur du pays le jour du recensement, y compris, par exemple, les fonctionnaires fédéraux et provinciaux en poste à l'extérieur du Canada, le personnel des ambassades canadiennes à l'étranger, les membres des Forces armées canadiennes en poste à l'étranger et les Canadiens qui se trouvent à bord de navires marchands. Comme le Recensement du Canada vise à dénombrer les personnes à l'endroit où elles résident habituellement plutôt qu'à l'endroit où elles se trouvent le jour du recensement, il est considéré comme un recensement *de jure*, ce qui signifie que les personnes se trouvant à l'étranger le jour du recensement ont été dénombrées si leur lieu habituel de résidence se trouvait au Canada. Certains pays mènent un recensement *de facto*, qui consiste à dénombrer les personnes à l'endroit où elles se trouvent le jour du recensement, même s'il ne s'agit pas de leur lieu habituel de résidence.

Au Recensement du Canada, divers questionnaires et formules sont utilisés pour recueillir les données. Les formules suivantes sont mentionnées dans le présent rapport.

La formule 1 est le Registre des visites (RV), qui sert à dresser la liste des logements privés et collectifs occupés et inoccupés ainsi que des exploitants et exploitations agricoles dans le secteur de dénombrement. Le RV est utilisé pour établir une liste d'adresses qui servira au cours des opérations de collecte et aux fins du contrôle.

La formule 2A est le questionnaire abrégé, qui comporte dix questions et est distribué à quatre ménages sur cinq. La formule 2B est un questionnaire complet qui renferme, outre les dix questions figurant sur la formule 2A, des questions permettant de recueillir des données sur divers sujets. Le questionnaire 2B est remis à un ménage sur cinq. Chaque ménage qui reçoit un questionnaire 2A ou 2B est invité à y dénombrer tous les membres du ménage qui sont inclus dans la population du recensement.

La formule 3 (A ou B) sert à dénombrer les personnes dans les logements collectifs (chaque résident d'un logement collectif doit remplir sa propre formule 3). Elle peut également être utilisée pour dénombrer des résidents habituels d'un ménage privé qui veulent remplir leur propre questionnaire du recensement plutôt qu'être inscrits sur un questionnaire 2A ou 2B. Une formule 4 est remplie par le personnel du recensement dans les cas où le ménage est absent ou que le logement est inoccupé le jour du recensement.

1.2 Méthodes de collecte

Pour assurer la meilleure couverture possible aux fins de la collecte, le Canada est divisé en petits secteurs géographiques appelés secteurs de dénombrement (SD). Chaque recenseur est affecté à un SD et remplit les tâches liées aux cartes, au listage, à la distribution et à la vérification; il s'assure en outre que tous les questionnaires sont retournés aux centres de dépouillement. Le nombre de ménages dans un SD varie de 175 en région rurale à 600 en région urbaine. Au recensement de 2001, on comptait 42 851 secteurs de dénombrement au Canada. Les recenseurs travaillent sous la surveillance de commissaires au recensement (CR). En 2001, les 2 917 CR affectés au recensement étaient chargés

d'embaucher les recenseurs et de veiller à la planification et à la gestion des activités de collecte sur le terrain dans leur zone d'affectation.

En 2001, environ 98 % des ménages ont été recensés selon la méthode de l'autodénombrement. Aux termes de cette méthode, un recenseur livre un questionnaire à chaque ménage au cours des deux semaines précédant le jour du recensement et un adulte ou un autre membre responsable du ménage est invité à remplir le questionnaire pour tous les membres du ménage, puis à le retourner dans une enveloppe-réponse.

Par ailleurs, environ 2 % des ménages ont été dénombrés par interview lors du recensement de 2001. Selon cette méthode, un recenseur visite le ménage et remplit lui-même un questionnaire pour le ménage par interview. Cette méthode est normalement utilisée dans les régions éloignées et nordiques du pays, de même que dans la plupart des réserves indiennes. Elle est en outre utilisée dans certaines régions urbaines où l'on estime qu'il existe une forte probabilité que les répondants ne retournent pas leur questionnaire.

Les recenseurs et les CR prennent part à un certain nombre d'activités de collecte sur le terrain. Ils doivent notamment communiquer avec des ménages afin de régler des problèmes qui ont souvent trait à l'exhaustivité ou à la cohérence de l'information fournie. Ils doivent aussi traiter les cas où les répondants n'ont pas retourné leur questionnaire.

Au cours des opérations de collecte sur le terrain, les recenseurs devaient livrer un questionnaire à chaque logement situé dans leur SD et inscrire l'adresse et le nom de la personne (si possible) dans leur Registre des visites (RV). Ils devaient en outre copier les numéros d'identification uniques devant être saisis ultérieurement et utilisés pour associer chaque ménage et chaque logement à la région géographique appropriée. Ensuite, ils devaient déterminer le numéro d'îlot correspondant au logement sur leur carte de SD et l'inscrire dans le RV et sur le questionnaire. Ces numéros d'îlot ont ensuite été saisis, de façon à permettre d'associer tous les logements du Canada à un îlot donné.

2. Traitement des données du recensement

2.1 Introduction

Cette étape du processus du recensement consiste à assurer le dépouillement de tous les questionnaires remplis, depuis la saisie des données jusqu'à la constitution d'une base de données d'extraction exacte et complète. Il s'agit ici de la saisie manuelle et automatisée des données, de l'imagerie des questionnaires, du contrôle, de la correction des erreurs, du codage, de l'imputation et de la pondération. La base de données finale a été transférée au Projet d'évaluation de la qualité des données, en vue de la détermination de la qualité globale des données, et au Projet de la diffusion, en vue de la production et de la commercialisation des produits et services du recensement de 2001. Dans le reste du présent chapitre, chaque opération de traitement des données est passée en revue.

Une innovation importante a été apportée au recensement de 2001; il s'agit de la création d'un système d'extraction d'images donnant accès aux images de tous les questionnaires et de tous les registres des visites (voir la section 2.3), de façon à éliminer la manutention de milliers de boîtes et de documents papier lors des processus subséquents nécessitant la consultation de l'original des formules du recensement, comme lors des recensements antérieurs.

2.2 Dépouillement régional

L'équipe du dépouillement régional était responsable de la saisie de l'information figurant dans les questionnaires en format électronique pour les systèmes de traitement ultérieurs ainsi que des recherches manuelles et du codage manuel des réponses aux questions sur l'industrie et à la profession tirées des questionnaires 2B. Étant donné l'énorme volume de questionnaires et d'information à saisir (plus de 4 milliards de frappes), la Tâche du dépouillement régional, depuis 1981, confie le travail en sous-traitance à l'Agence des douanes et du revenu du Canada (ADRC), anciennement Revenu Canada. Celle-ci utilise son réseau de systèmes, de ressources et de personnel pour saisir et coder les données du recensement. Ce partenariat permet à l'équipe du recensement de réaliser des économies en mettant à profit le personnel qualifié et l'infrastructure déjà en place à l'ADRC. En ce qui a trait à la qualité des données, il lui permet en outre de bénéficier de l'expérience acquise par l'ADRC dans le cadre du dépouillement des questionnaires des recensements antérieurs. Lors du recensement de 2001, environ 2 800 employés de l'ADRC ont prêté le serment d'office et de discrétion prévu par la *Loi sur la statistique* en vue d'effectuer les tâches du recensement. Aux termes de cette loi, les employés de l'ADRC sont soumis aux mêmes règles et règlements que les employés de Statistique Canada.

Une fois les activités de collecte relatives à un secteur de dénombrement (SD) terminées, les questionnaires, les cartes géographiques et les registres des visites connexes étaient acheminés dans des boîtes de SD des sous-sections de la collecte sur le terrain vers un des huit centres fiscaux de l'ADRC désignés du pays.

La première étape consistait à préparer les questionnaires remplis en vue de la saisie des données. Traditionnellement, elle avait pour objet d'attribuer manuellement des codes aux réponses en lettres fournies par les répondants. En 2001, la plupart des réponses en lettres étaient converties en codes par des systèmes automatisés (voir la section 2.5). Les seules réponses en lettres devant être codées manuellement lors du recensement de 2001 étaient celles du questionnaire complet (2B) portant sur les questions relatives à l'industrie et la profession. Des travaux de recherche ont été entrepris en vue d'automatiser le codage de ces réponses et on s'attend à disposer d'un système automatisé opérationnel pour le recensement de 2006.

Les réponses aux questions sur l'industrie ont été codées à l'ADRC conformément au Système de classification des industries de l'Amérique du Nord (SCIAN), que Statistique Canada utilise comme document normatif depuis quelques années. Le SCIAN a été élaboré afin de fournir au Canada, aux États-Unis et au Mexique un cadre de référence commun pour la production de statistiques sur l'industrie

aux termes de l'Accord de libre-échange nord-américain (ALENA). Cela représente une modification au codage des réponses relatives à l'industrie qui, lors du recensement de 1996, avaient été codées à l'aide de la Classification type des industries (CTI) de 1980. Afin de permettre l'établissement de comparaisons longitudinales, les réponses recueillies sur les questionnaires complets (2B) à la question sur l'industrie du recensement de 2001 sont également codées à l'aide de la CTI de 1980 dans le cadre du codage automatisé (voir la section 2.5).

Une fois que les questionnaires avaient été reçus et enregistrés à un des centres fiscaux de l'ADRC et que les codes d'industrie et de profession avaient été attribués, l'étape suivante consistait à trier, étiqueter et mettre en lots les questionnaires en vue de la saisie des données. L'étiquette apposée sur chacun des questionnaires portait un numéro de séquence unique destiné à permettre de contrôler le cheminement du questionnaire tout au long des opérations à l'ADRC. Pour la première fois, l'étiquette comportait aussi un code à barres ayant pour objet de faciliter le balayage du questionnaire lors de la Tâche de l'imagerie (voir la section 2.3).

La saisie des données se faisait ensuite selon la méthode traditionnelle de l'entrée au clavier. Afin de vérifier la précision de l'opération de saisie, on saisissait de nouveau les données d'un échantillon de questionnaires dont les données avaient déjà été saisies. Des statistiques sur le contrôle de la qualité étaient produites en comparant entre eux les deux ensembles de données saisies. Comme on peut s'y attendre, l'entrée des données au clavier à partir des questionnaires peut se traduire par des erreurs, attribuables à diverses raisons : erreur de saisie, réponses difficiles à lire ou mal indiquées sur les questionnaires ou réponses omises par l'opérateur de saisie des données. La vérification des entrées permet de réduire les erreurs le plus possible.

Au fur et à mesure de leur saisie, les données étaient transmises en temps réel sur des lignes de communication dédiées à l'ordinateur de l'ADRC à Ottawa. À l'intérieur d'un délai de 24 heures, les données étaient ensuite transférées sur des cartouches magnétiques qui étaient livrées par un transporteur cautionné à Statistique Canada, où elles étaient chargées dans l'ordinateur central. Les questionnaires étaient enfin rassemblés dans leurs boîtes de SD en vue de leur expédition au centre de dépouillement de Statistique Canada à Ottawa. Une fois que les données avaient été saisies et transférées à Statistique Canada et que celui-ci avait accusé réception de toutes les données, l'ADRC n'avait plus aucune donnée du recensement en sa possession.

2.3 Imagerie

Lors des recensements antérieurs, les étapes du dépouillement qui nécessitaient la consultation des questionnaires et des registres des visites (RV) se faisaient à l'aide des documents papier. En 2001, on a éliminé la nécessité de manipuler les documents papier en créant (par balayage) une image électronique de tous les questionnaires et RV dès leur réception au centre de dépouillement du recensement en provenance des centres fiscaux de l'Agence des douanes et du revenu du Canada (ADRC). Le personnel des opérations ultérieures pouvait ensuite accéder aux images des questionnaires et des RV à l'aide d'un système d'extraction d'images, plutôt que d'avoir à utiliser les documents papier.

À leur arrivée au centre de dépouillement du recensement de 2001, les boîtes de secteurs de dénombrement (SD) étaient enregistrées et les documents étaient préparés en vue du balayage.

Les images électroniques des 13 millions de documents (en majorité des questionnaires) ont été créées à l'aide de 15 scanners à haute vitesse fonctionnant cinq jours par semaine, pendant deux quarts de travail par jour. L'identificateur géographique nécessaire pour identifier chacune des images de document était automatiquement attribué à partir du code à barres figurant sur l'étiquette apposée lors des opérations de saisie des données à l'ADRC (voir la section 2.2). Un contrôle qualitatif était effectué afin de vérifier que chaque document comportait le nombre approprié de pages et que le nombre de questionnaires de chaque type était exact pour chacun des SD. Une opération spéciale était prévue afin de résoudre les problèmes qui se posaient. Les images étaient alors enregistrées sur des disques optiques à des fins de consultation ultérieure et d'archivage. Elles étaient aussi conservées sur une

mémoire cache magnétique afin qu'on puisse y avoir accès immédiatement pour les besoins des activités de la Tâche de la vérification interactive.

2.4 Vérification interactive

L'objectif principal de la Vérification interactive était d'identifier et de corriger les erreurs dans les données dont la résolution exigeait de se référer aux images des questionnaires et/ou des registres des visites. Les données saisies étaient soumises à un ensemble détaillé de règles de vérification pour identifier des erreurs telles que des personnes manquantes ou comptées plus d'une fois dans un ménage; un dénombrement incorrect des résidents étrangers ou temporaires; des ménages associés aux mauvais questionnaires; ou une mauvaise classification des logements (occupés ou inoccupés). Un examen exhaustif de l'information des formules associées à chacun des problèmes était effectué pour déterminer la correction appropriée. Dans certains cas il fallait ajouter et/ou éliminer des personnes ou des logements.

Lorsque les cartouches renfermant les données étaient reçues de l'Agence des douanes et du revenu du Canada (ADRC), elles étaient chargées sur les ordinateurs de Statistique Canada en vue de la Vérification interactive. Une série de contrôles structurels automatisés étaient alors exécutés, surtout pour vérifier les renseignements inscrits par le recenseur sur la page couverture des questionnaires. Ces contrôles avaient entre autres pour objet d'apparier les types de questionnaire et les types de ménage, de contre-vérifier le nombre de questionnaires, de contre-vérifier le nombre de personnes dénombrées et de vérifier si les identificateurs géographiques étaient bien uniques. Les données sur le revenu tirées des questionnaires 2B étaient aussi soumises à certains contrôles visant à déceler les anomalies et à les soumettre aux spécialistes des données sur le revenu.

Tous les contrôles étaient effectués par secteur de dénombrement (SD). Les erreurs étaient balisées, puis corrigées en se reportant aux images des questionnaires et des registres des visites (RV) du SD visé. Les corrections apportées aux données électroniques étaient faites à l'aide d'un système interactif sur micro-ordinateur. Certaines corrections étaient également notées sur les images de questionnaire ou sur les registres de visite à l'aide d'un processus électronique.

Une fois que la boîte de SD avait été soumise à tous ces contrôles, on procédait à une vérification manuelle et automatique des identificateurs géographiques transcrits par le recenseur de la carte de SD au questionnaire et au RV.

La Vérification interactive prévoyait aussi un dépouillement spécial afin d'assurer que les Canadiens qui résidaient à l'extérieur du Canada le jour du recensement (personnes à bord des navires de la garde côtière, des navires de guerre et des navires marchands battant pavillon canadien, membres du personnel diplomatique et membres des Forces armées) avaient été dénombrés correctement.

La dernière étape de la Vérification interactive consistait à reformater les données et à les acheminer aux dernières étapes du dépouillement, c.-à-d. le Codage automatisé ainsi que le Contrôle et imputation.

2.5 Codage automatisé et codage interactif

Le codage automatisé consistait à convertir en codes les réponses en lettres saisies sur les questionnaires complets lors du dépouillement régional (voir la section 2.2), à l'aide d'un fichier de référence ou d'une structure de classification automatisé contenant des séries de mots ou d'expressions et leur code numérique correspondant. Bien que la majorité des réponses en lettres puissent être codées de façon entièrement automatique, certaines réponses ne peuvent être appariées. Des systèmes de codage interactifs pour PC sont prévus pour aider des codeurs ayant reçu une formation spéciale et des spécialistes à attribuer un code approprié à toutes les réponses non appariées en examinant les réponses données à d'autres questions, lesquelles se rapportaient parfois à d'autres membres du

ménage. Ont été soumises au codage automatisé les réponses en lettres aux questions suivantes du questionnaire complet (2B) :

- lien avec la Personne 1;
- langue parlée à la maison;
- langues non officielles;
- première langue apprise dans l'enfance (langue maternelle);
- langue utilisée au travail;
- lieu de naissance;
- lieu de naissance des parents;
- citoyenneté;
- origine ethnique (ascendance);
- groupe de population;
- bande indienne/Première nation;
- lieu de résidence 1 an auparavant;
- lieu de résidence 5 ans auparavant;
- principal domaine d'études;
- religion;
- lieu de travail; et
- industrie selon le Système de classification des industries de l'Amérique du Nord (et plus tard selon la Classification type des industries de 1980).

Une fois codées toutes les réponses associées à une variable, les données étaient acheminées à l'étape du Contrôle et imputation.

2.6 Contrôle et imputation

Les données recueillies lors de toute enquête comportent des omissions ou des incohérences. Ces erreurs peuvent survenir lorsque les répondants fournissent des réponses incorrectes ou incomplètes aux questions ou encore lors du traitement. Par exemple, un répondant peut être réticent à répondre à une question, peut ne pas se souvenir de la bonne réponse ou peut mal comprendre la question. De son côté, le personnel du recensement peut coder les réponses incorrectement ou peut faire d'autres types d'erreurs au cours du traitement.

Une des premières tâches de l'opération de contrôle et d'imputation consiste à vérifier si la taille du ménage est indiquée pour tous les logements classés comme étant « occupés ». Pour les logements occupés associés non pas à un questionnaire ordinaire (formule 2A ou 2B) mais à une formule 4 établie en cas de non-réponse, la première mesure prise à l'opération du contrôle et de l'imputation était de s'assurer que la taille du ménage était valide. Si la taille du ménage était « inconnue », la procédure consistait à imputer la taille du ménage à partir du plus proche voisin. Cependant, une nouvelle procédure a été introduite en 2001 afin d'estimer la taille du ménage dans le cas des logements associés à une formule 4. Pour de plus amples renseignements, se reporter au chapitre 7 portant sur l'Étude sur la classification des logements.

L'épuration définitive des données, faite à l'étape du contrôle et de l'imputation, a été presque entièrement automatisée. Elle consistait à appliquer une série de règles de contrôle détaillées, qui permettaient de repérer les réponses manquantes ou incohérentes. Ces dernières étaient la plupart du temps corrigées en modifiant par imputation la valeur du plus petit nombre de variables possible. L'imputation était faite à l'aide des méthodes dites « déterministe » et/ou « hot deck à changements minimaux ». L'imputation déterministe consiste à corriger les erreurs en inférant la valeur appropriée à

partir des réponses à d'autres questions. Quant à l'imputation « hot deck à changements minimaux », on choisit un enregistrement qui comporte un certain nombre de caractéristiques communes à celles de l'enregistrement qui contient une erreur. On emprunte les données de cet enregistrement « donneur » et on les utilise pour changer le plus petit nombre possible de variables nécessaires pour résoudre tous les rejets au contrôle.

Deux systèmes automatisés différents ont été utilisés pour mener à bien ce traitement.

Le système NIM (**N**earest-**n**eighbour **I**mputation **M**ethod), qui a été élaboré en vue du recensement de 1996 pour faire le contrôle et l'imputation des caractéristiques démographiques de base telles que l'âge, le sexe, l'état matrimonial, l'union libre et le lien avec la Personne 1, a été amélioré en vue du recensement de 2001 et implanté dans un système appelé SCANCIR (**S**ystème **c**anadien de **c**ontrôle et d'**i**mputation du recensement). Il permet maintenant d'effectuer le contrôle et l'imputation d'autres variables telles que l'industrie, le lieu de travail, le mode de transport et la mobilité. Tout comme en 1996, le système SCANCIR permettait de soumettre les réponses à des contrôles plus détaillés et exacts, tout en réservant les réponses grâce à la méthode d'imputation « hot deck à changements minimaux ».

Le système SPIDER (**S**ystem for **P**rocessing **I**nstructions from **D**irectly **E**ntered **R**equirements), qui a été utilisé pour traiter les autres variables du recensement telles que la langue maternelle, le logement ou le revenu, permet de traduire les exigences des agents spécialisés, déterminées au moyen de tables de décision logique, en modules exécutables. SPIDER fait appel aux imputations déterministe et « hot deck ».

2.7 Ajustements de la couverture en fonction des logements inoccupés et des logements non répondants

L'Étude sur la classification des logements (ÉCL) consiste à prendre un échantillon de logements ayant été inscrits comme étant soit inoccupés ou occupés lors de la collecte. Plus tard, l'ÉCL retourne à ces logements pour déterminer s'ils étaient occupés ou non le jour du recensement, ou s'ils n'auraient pas dû être inscrits parce qu'ils ne correspondaient pas à la définition d'un logement aux fins du recensement.

S'il s'avère qu'un logement était occupé, on effectue un de deux ajustements possibles dans la base de données du recensement. Pour les logements inscrits comme étant inoccupés le jour du recensement, on applique une technique d'estimation faisant appel à des « additions aléatoires », afin d'ajouter des logements et des personnes à la base de données du recensement. Au recensement de 2001, 111 628 ménages et 222 720 personnes ont ainsi été ajoutés à la base de données pour tenir compte du nombre estimatif de personnes vivant dans des logements « inoccupés ». Le deuxième ajustement possible a trait aux logements occupés pour lesquels aucun questionnaire du recensement rempli n'a été reçu (c'est-à-dire les logements non répondants) et consiste à créer, dans la base de données du recensement, une nouvelle taille de ménage pour tous ces logements. Au total, 143 681 ménages comptant 317 587 personnes ont été ajoutés à la base de données par suite de cet ajustement.

L'ÉCL est la seule étude sur la couverture qui se traduit par l'ajout de ménages et de personnes à la base de données du recensement. En revanche, les trois autres études sur la couverture décrites dans le présent rapport n'entraînent aucun changement à la base de données. Les résultats de ces études influent sur le calcul des chiffres de population officiels de Statistique Canada.

Le chapitre 7 renferme des renseignements plus détaillés sur la portée et la méthodologie de l'ÉCL.

2.8 Pondération

On a recueilli des renseignements sur l'âge, le sexe, l'état matrimonial, l'union libre, la langue maternelle et le lien avec la Personne 1 auprès de presque toute la population. Cependant, la majeure partie des renseignements proviennent d'un échantillon de 20 % de la population (un ménage sur cinq) à qui on a

livré un questionnaire complet (2B) (voir la section 1.1). On a employé la pondération pour reporter, après le contrôle et l'imputation, l'information fournie par l'échantillon à l'ensemble de la population.

La méthode de pondération a fourni des poids qui, à partir des données-échantillon, servaient à produire des estimations. Au recensement de 2001, on a fait appel à une technique que l'on appelle estimation par calibrage ou par régression. Cette procédure consiste à ajuster des poids initiaux égalant environ 5 par le plus petit nombre qui permette de faire correspondre d'aussi près que possible les estimations échantillonnelles (p. ex., nombre d'hommes ou de personnes âgées de 15 à 19 ans) aux chiffres de population obtenus à partir des questionnaires abrégés (2A) pour l'âge, le sexe, l'état matrimonial, l'union libre et la taille du ménage.

3. Évaluation de la qualité des données

3.1 Généralités

Tout au long des différentes étapes du recensement, on s'est efforcé d'obtenir des résultats de qualité supérieure. Des normes de qualité rigoureuses ont été établies pour la collecte et le traitement des données; par ailleurs, le Programme d'information publique a contribué à réduire les cas de non-réponse. Un Programme d'évaluation de la qualité des données a également été mis sur pied afin de fournir aux utilisateurs de l'information sur la qualité des données du recensement et sur leurs limites.

Malgré les efforts considérables déployés au cours du processus pour obtenir des données de bonne qualité, les données peuvent présenter certaines inexactitudes. Pour évaluer l'utilité des données du recensement dans leurs applications et comprendre les risques qu'il y a à tirer des conclusions ou à prendre des décisions en se fondant sur ces données, les utilisateurs doivent être au fait de leurs inexactitudes et savoir quelles sont leur origine et leur composition.

Le présent rapport porte sur les questions relatives à la couverture du recensement de 2001 et présente des données tirées de diverses enquêtes postcensitaires visant à mesurer le sous-dénombrement et le surdénombrement. Parmi la série des Rapports techniques du recensement de 2001, les utilisateurs trouveront de l'information détaillée supplémentaire sur le recensement de 2001 dans le rapport sur *l'Échantillonnage et la pondération* (n° 92-395-XIF au catalogue).

Dans tout recensement, il est impossible d'atteindre une couverture parfaite, et ce, en dépit des suivis et des vérifications faites sur le terrain par les recenseurs et les commissaires au recensement, et des mesures décrites à la section 3.2. C'est pourquoi on procède à des études sur la couverture afin de mesurer l'ampleur des erreurs de couverture qui se produisent lorsque des logements ou des personnes sont omis, dénombrés incorrectement ou comptés deux fois. Dans le présent rapport, il sera question de l'Étude sur la classification des logements (chapitre 7), de la Contre-vérification des dossiers (chapitre 8), de l'Étude par appariement automatisé (chapitre 9) et de l'Étude sur les logements collectifs (chapitre 10).

3.2 Causes des erreurs de couverture et mesures prises en vue d'en réduire le nombre

Dans la plupart des cas, les erreurs de couverture surviennent durant l'étape de la collecte. Par exemple, le sous-dénombrement des personnes et des ménages se produit lorsque des logements occupés sont omis complètement ou qu'ils sont classés à tort comme logements inoccupés. Il y a aussi sous-dénombrement de la population si une personne est omise au sein d'un ménage dénombré. Le surdénombrement peut survenir lorsqu'il y a incertitude quant au lieu habituel de résidence d'une personne, par exemple, dans le cas d'un étudiant de niveau universitaire qui est dénombré au domicile d'un parent et aussi à la résidence d'étudiants.

Des erreurs de couverture peuvent également se glisser au cours du dépouillement lorsque des enregistrements de personnes ou de ménages sont supprimés par erreur, égarés ou créés artificiellement.

Ces sources d'erreurs potentielles ont été déterminées au cours de la planification du recensement de 2001, et certaines mesures de contrôle ont été prises en vue d'en réduire le nombre. Ces mesures comprennent :

- a) la délimitation et cartographie précise des secteurs de dénombrement (SD) afin de s'assurer qu'aucune aire n'ait été omise ou incluse deux fois;

- b) les instructions dans le manuel du recenseur sur la façon de faire la prospection de son SD afin de réduire les risques d'omission de logements;
- c) la création d'un Registre des adresses à partir de sources autres que le recensement et l'utilisation de cette liste par le recenseur pour vérifier si des logements ont été omis;
- d) la localisation préalable des logements collectifs que le personnel sur le terrain devait vérifier pour s'assurer qu'ils seraient dénombrés au recensement s'ils étaient occupés;
- e) les procédures spéciales pour dénombrer les personnes qui vivent dans certains secteurs spéciaux du centre des principales villes et qui ont de la difficulté à répondre (p. ex., difficulté en français ou en anglais, ou problèmes de littératie);
- f) les procédures spéciales pour dénombrer la population vivant dans les réserves ou établissements indiens;
- g) les messages publicitaires pour informer les Canadiens sur le recensement, qui indiquaient les mesures à prendre s'ils n'avaient pas reçu de questionnaire;
- h) les instructions sur les personnes à inscrire dans le questionnaire du recensement pour rappeler aux répondants quelles personnes devaient être incluses;
- i) les questions dans le questionnaire du recensement visant à déterminer si le répondant avait des doutes quant à l'inscription de certaines personnes et le suivi dans ces cas pour aider le répondant.

Ces mesures, ajoutées à la formation appropriée, aux vérifications effectuées par les surveillants et aux systèmes de contrôle qualitatif appliqués au cours des étapes de collecte et de dépouillement du recensement, ont contribué à réduire le nombre d'erreurs de couverture. Elles n'ont toutefois pas permis de les éliminer complètement, d'où l'importance d'évaluer l'ampleur des erreurs de couverture.

Au recensement de 2001, une situation spécifique touchant la couverture s'est produite dont les utilisateurs des données du recensement doivent être au fait. Dans certaines réserves indiennes et certains établissements indiens, le dénombrement a dû être interrompu avant d'être mené à terme. Par ailleurs, certaines réserves indiennes et certains établissements indiens ont été dénombrés trop tard pour pouvoir être inclus ou, encore, la qualité des données recueillies a été jugée inadéquate. Un problème similaire a été constaté aux recensements de 1996, de 1991 et de 1986 et, dans une moindre mesure, au recensement de 1981. Ces réserves indiennes et établissements indiens (30 en tout en 2001) sont appelés réserves indiennes et établissements indiens partiellement dénombrés. Les données de 2001 ne sont donc pas disponibles pour ces secteurs et n'ont pas été incluses dans les produits du recensement de 2001. Même si ces réserves et établissements indiens ne sont pas inclus dans les produits du recensement de 2001, on en produit les estimations de population fondées sur un modèle statistique. Cette question est traitée au chapitre 14 du présent rapport. Pour obtenir de l'information supplémentaire concernant les réserves et établissements indiens, voir Dick et So (2004).

3.3 Erreurs de couverture – Définitions

Les erreurs de couverture sont des erreurs qui ont une incidence sur l'exactitude des chiffres des divers univers du recensement (voir le chapitre 4). Il existe deux genres d'erreurs de couverture : les erreurs de sous-dénombrement et les erreurs de surdénombrement. Il y a sous-dénombrement lorsqu'on omet complètement de dénombrer une unité faisant partie d'un univers visé par le recensement. Par ailleurs, le surdénombrement peut survenir de deux façons. La première, et la plus fréquente, est lorsqu'une unité faisant partie d'un univers du recensement est dénombrée plus d'une fois. La seconde est lorsqu'une unité ne faisant pas partie de l'univers du recensement (par exemple un résident étranger, une personne fictive ou un logement marginal inoccupé) est dénombrée par erreur. Il faut prendre note qu'une erreur

d'ordre géographique uniquement ne constitue pas une erreur de couverture. Par exemple, une personne qui a été dénombrée dans la mauvaise région géographique ne constitue ni un cas de surdénombrement dans la région où elle a été dénombrée ni un cas de sous-dénombrement dans la région où elle aurait dû l'être.

Le sous-dénombrement d'un ménage correspond au cas d'un ménage dont tous les membres sont omis. Les ménages dont certains membres mais pas tous sont omis ne constituent pas des cas de sous-dénombrement de ménages, même s'ils entraînent des erreurs dans les données sur les caractéristiques du ménage, comme sa taille et sa composition. De même, le surdénombrement d'un ménage se produit lorsque tous les membres sont dénombrés plus d'une fois.

Le sous-dénombrement, le surdénombrement et la valeur nette des deux, le sous-dénombrement net, sont le plus utilement exprimés en taux de population qui aurait dû être dénombrée lors du recensement. En termes mathématiques, supposons que U représente le nombre total ou « réel » d'unités dans l'univers visé et que R corresponde aux chiffres officiels du recensement pour cet univers. Étant donné que le concept est le même pour les personnes, les ménages, les logements et les familles, on emploiera alors le terme général « unités ». Dans ce cas, l'erreur qui résulte de l'utilisation de R au lieu de U est :

$$N = U - R$$

Cette erreur, désignée par N , s'appelle l'**erreur de couverture nette**.

Supposons que O représente le sous-dénombrement total, c'est-à-dire le nombre total d'unités dans l'univers en question qui ont été omises lors du recensement, et que D corresponde au nombre total d'unités dans l'univers qui ont été dénombrées au moins une fois. Dans ce cas,

$$U = O + D$$

et par conséquent,

$$\begin{aligned} N &= (O + D) - R \\ &= O - (R - D) \\ &= O - S \end{aligned}$$

où le terme $S = R - D$ correspond au surdénombrement. Cette erreur découle non seulement du dénombrement répété des mêmes unités, mais aussi du dénombrement d'unités qui ne font pas partie de l'univers visé.

Exprimés en proportion du nombre d'unités dans l'univers visé, les taux d'erreur de couverture sont :

- a) **taux de sous-dénombrement** : $T_O = O/U$;
- b) **taux de surdénombrement** : $T_S = S/U$;
- c) **taux de sous-dénombrement net** : $T_N = N/U$;

de sorte que

$$T_N = (O - S)/U = T_O - T_S$$

Un taux de sous-dénombrement net positif signifie que le sous-dénombrement est supérieur au surdénombrement, alors qu'un taux de sous-dénombrement net négatif indique que le sous-dénombrement est inférieur au surdénombrement. Comme dans la plupart des cas le sous-dénombrement est supérieur au surdénombrement, la plupart des chiffres et des taux de sous-dénombrement net sont positifs.

4. Univers du recensement

4.1 Introduction

Les cinq univers suivants ont été dénombrés au recensement de 2001 :

1. l'univers de la population;
2. l'univers des logements;
3. l'univers des ménages;
4. l'univers des familles de recensement;
5. l'univers des familles économiques.

Chacun de ces univers est décrit ci-après, mais il convient de noter que le Programme de mesure de l'erreur de couverture de 2001 a porté principalement sur les univers de la population et des logements. Les lecteurs sont invités à consulter le *Dictionnaire du recensement de 2001* (n° 92-378-XIF au catalogue), qui contient plus de détails sur les variables associées à chacun des univers.

4.2 Univers de la population

L'univers de la population du recensement de 2001 comprend les groupes de personnes suivants :

- les citoyens canadiens et les immigrants reçus ayant un lieu habituel de résidence au Canada;
- les citoyens canadiens et les immigrants reçus qui sont à l'extérieur du Canada comme employés du gouvernement canadien (fédéral, provincial et territorial) ou membres des Forces armées canadiennes, ainsi que leur famille;
- les citoyens canadiens et les immigrants reçus qui sont en mer ou dans des ports à bord de navires marchands battant pavillon canadien;
- les résidents non permanents :
 - les personnes au Canada qui demandent le statut de réfugié;
 - les personnes au Canada qui détiennent un permis de séjour pour étudiants (les étudiants étrangers, les titulaires d'un visa pour étudiants);
 - les personnes au Canada qui détiennent un permis de travail (les travailleurs étrangers, les titulaires d'un permis de travail);
 - les personnes au Canada qui détiennent un permis ministériel (y compris les prolongations);
 - toutes les personnes nées à l'extérieur du Canada et qui sont à la charge de personnes revendiquant le statut de réfugié ou de personnes titulaires d'un permis de séjour pour étudiants, d'un permis de travail ou d'un permis ministériel.

Les groupes de personnes suivants, qui, ensemble, constituent la catégorie des résidents étrangers, n'ont pas été inclus dans l'univers de la population du recensement de 2001 :

- les représentants du gouvernement d'un autre pays qui sont affectés à l'ambassade, au haut-commissariat ou à toute autre mission diplomatique de ce pays au Canada, et leur famille;
- les membres des forces armées d'un autre pays en poste au Canada, et leur famille;
- les résidents d'un autre pays en visite au Canada temporairement.

Depuis le recensement de 1991, les résidents non permanents sont inclus dans l'univers de la population. Auparavant, les résidents non permanents étaient considérés comme des résidents étrangers et n'étaient pas inclus dans l'univers de la population. Les utilisateurs doivent tenir compte de ce fait lorsqu'ils

établissent des comparaisons entre les données des recensements de 1991, de 1996 et de 2001 et celles des recensements antérieurs.

Les définitions qui figurent ci-dessus indiquent quelles personnes doivent être incluses dans le recensement, mais elles ne font pas état de l'endroit où ces personnes doivent être dénombrées. Comme on l'a déjà mentionné, le Recensement du Canada fait appel à la méthode de dénombrement *de jure*, selon laquelle les personnes doivent être dénombrées à leur lieu habituel de résidence même si elles en sont temporairement absentes au moment de la tenue du recensement. Les personnes qui sont absentes de leur lieu habituel de résidence au moment du recensement et qui résident ailleurs au Canada doivent être dénombrées à leur lieu habituel de résidence et être considérées comme des « résidents temporaires » à l'autre endroit (à ne pas confondre « résidents temporaires » avec « résidents non permanents », concept qui se rapporte au statut légal de la personne pendant son séjour au Canada). Les personnes qui n'ont pas de lieu habituel de résidence doivent être dénombrées là où elles se trouvent le jour du recensement. D'autres pays utilisent la méthode *de facto*, selon laquelle les personnes doivent être dénombrées à l'endroit où elles se trouvent le jour du recensement quel que soit leur lieu habituel de résidence.

4.3 Univers des logements

Un logement est un ensemble de pièces d'habitation qu'une personne ou un groupe de personnes habitent ou pourraient habiter. Seuls les logements situés au Canada sont inclus dans l'univers des logements. Il existe deux types de logement :

- Un logement privé est un ensemble distinct de pièces d'habitation ayant une entrée privée donnant sur l'extérieur ou sur un corridor, un hall, un vestibule ou un escalier commun à l'intérieur. L'entrée doit donner accès au logement sans que l'on ait à passer par les pièces d'habitation de quelqu'un d'autre.
- Un logement collectif est un établissement commercial, institutionnel ou communautaire que le recenseur peut identifier comme tel grâce à une enseigne ou en s'adressant à la personne qui en a la charge, à un résident, à un voisin, etc. Sont inclus dans cette catégorie les pensions et maisons de chambre, les hôtels, motels et maisons de chambres pour touristes, les maisons de repos, les hôpitaux, les résidences de personnel, les casernes, les camps de chantier, les prisons, les centres d'accueil, les foyers collectifs, etc.

Ces deux types de logements peuvent être classés dans des catégories plus détaillées :

- Les logements privés peuvent être classés comme logements privés ordinaires, logements marginaux ou logements en construction. Les logements privés ordinaires se subdivisent en trois catégories : logements occupés par des résidents habituels, logements occupés par des résidents étrangers ou temporaires seulement et les logements inoccupés. Les logements marginaux et les logements en construction sont classés comme logements occupés par des résidents habituels ou comme logements occupés par des résidents étrangers ou temporaires seulement. Les logements marginaux et les logements en construction inoccupés le jour du recensement ne sont pas inclus dans l'univers des logements.
- Les logements collectifs sont classés comme logements occupés par des résidents habituels, logements occupés par des résidents étrangers ou temporaires seulement et logements collectifs inoccupés. Dans le cas des logements collectifs inoccupés, des données ont été recueillies, mais ne sont pas incluses dans les produits du recensement.

En résumé, l'univers des logements inclut :

- les logements privés ordinaires occupés par des résidents habituels;
- les logements privés ordinaires occupés par des résidents étrangers ou temporaires seulement;
- les logements privés ordinaires inoccupés;
- les logements marginaux et les logements en construction en autant qu'ils étaient occupés le jour du recensement;
- les logements collectifs occupés par des résidents habituels;
- les logements collectifs occupés par des résidents étrangers ou temporaires seulement.

L'univers des logements n'inclut pas :

- les logements marginaux et les logements en construction inoccupés le jour du recensement;
- les logements collectifs inoccupés le jour du recensement;
- les logements situés à l'extérieur du Canada.

4.4 Univers des ménages

Un ménage est une personne ou un groupe de personnes, autres que les résidents étrangers ou temporaires, occupant un même logement et n'ayant pas de lieu habituel de résidence ailleurs au Canada. Il peut se composer d'un groupe familial (famille de recensement) avec ou sans autres personnes hors famille de recensement, de deux familles ou plus partageant le même logement, d'un groupe de personnes non apparentées ou d'une personne vivant seule. Les membres du ménage temporairement absents le jour du recensement sont considérés être membres du ménage à leur lieu habituel de résidence. Aux fins du recensement, chaque personne est membre d'un seul et unique ménage.

Les ménages sont classés en trois genres selon le type de logement dans lequel ils résident : ménages privés, ménages collectifs et ménages à l'extérieur du Canada. Il faut prendre note que les ménages à l'extérieur du Canada ne sont pas associés à un logement, car les logements situés à l'extérieur du Canada ne sont pas compris dans l'univers des logements. La plupart des données du recensement diffusées sur les ménages portent sur les ménages privés seulement.

4.5 Univers des familles de recensement

Une famille de recensement est :

- un couple actuellement marié avec ou sans enfants des deux conjoints ou de l'un d'eux habitant dans un même logement;
- un couple vivant en union libre avec ou sans enfants des deux partenaires ou de l'un d'eux;
- un parent seul, quel que soit son état matrimonial, demeurant avec au moins un enfant dans le même logement.

Les données sur les familles de recensement ne sont recueillies que pour les personnes dans les ménages privés et les ménages à l'extérieur du Canada.

4.6 Univers des familles économiques

Une famille économique est un groupe de deux personnes ou plus qui vivent dans le même logement et qui sont apparentées par le sang, par alliance, par union libre ou par adoption. Les données sur les familles économiques ne sont recueillies que pour les personnes dans les ménages privés.

4.7 Liens existant entre les univers

Le tableau 4.7 présente une récapitulation des trois univers de base selon l'emplacement du logement (au Canada ou à l'extérieur du Canada), le classement du logement (privé ou collectif) et son occupation/inoccupation. On indique si chaque groupe dans le tableau est inclus dans l'univers ou s'il en est exclu. Par exemple, les employés du gouvernement canadien vivant à l'étranger avec leur famille sont inclus dans les univers de la population et des ménages, mais ils sont exclus de l'univers des logements. Les catégories signalées par un astérisque (*) correspondent à des groupes pour lesquels les données sont recueillies mais qui sont exclus de la plupart des produits du recensement.

Tableau 4.7 Les univers de la population, des ménages et des logements et les liens existant entre eux

	Population	Ménages	Logements
Logements au Canada			
1. Logements privés – ordinaires			
• occupés par des résidents habituels	I	I	I
• occupés par des résidents étrangers ou temporaires seulement	E	E	I*
• inoccupés	–	–	I*
2. Logements privés – marginaux ou en construction			
• occupés par des résidents habituels	I	I	I
• occupés par des résidents étrangers ou temporaires seulement	E	E	I*
• inoccupés	–	–	E
3. Logements collectifs			
• occupés par des résidents habituels	I	I*	I*
• occupés par des résidents étrangers ou temporaires seulement	E	E	I*
• inoccupés	–	E	E
4. Additions aléatoires à la suite de la classification des logements			
	I	I	I
Logements à l'extérieur du Canada			
5. Citoyens canadiens et immigrants reçus qui sont à l'étranger, dans une base militaire ou en mission diplomatique			
	I	I*	E
6. Citoyens canadiens et immigrants reçus qui sont en mer ou dans des ports à bord de navires marchands battant pavillon canadien			
	I	I*	E

I = inclus, E = exclus, – = sans objet

* Indique que les données ont été recueillies, mais ne sont pas incluses dans la plupart des produits.

5. Sous-dénombrement net au recensement de 2001

Au recensement de 2001, le taux de sous-dénombrement net de la population a atteint 2,99 %. Cela signifie que selon les résultats nets, 2,99 % des personnes (924 429 personnes) qui auraient dû être dénombrées au recensement ont été omises. Le taux de sous-dénombrement brut de la population de 3,95 % a été contrebalancé par le taux de surdénombrement de 0,96 %. Ce dernier taux correspond au pourcentage de personnes dénombrées par erreur au recensement alors que le premier indique que 3,95 % de la population visée du recensement n'a pas été dénombrée.

Dans ce chapitre, nous présentons des estimations du sous-dénombrement net selon diverses caractéristiques du recensement. Le tableau 5.1 montre le sous-dénombrement net estimé et l'erreur-type estimée ainsi que le taux de sous-dénombrement net et l'erreur-type correspondante pour chacune des caractéristiques qui y figurent. Les estimations fournies se rapportent aux personnes.

Tableau 5.1 Estimations du sous-dénombrement net de la population et de l'erreur-type au recensement de 2001, selon diverses caractéristiques, Contre-vérification des dossiers de 2001

Caractéristiques	Nombre net de personnes omises		Taux de sous-dénombrement net de la population	
	Nombre estimé (%)	Erreur-type	Taux estimé (%)	Erreur-type
Canada	924 429	44 749	2,99	0,14
Terre-Neuve-et-Labrador	9 401	1 782	1,80	0,33
Île-du-Prince-Édouard	1 325	775	0,97	0,56
Nouvelle-Écosse	24 521	4 170	2,63	0,44
Nouveau-Brunswick	20 095	3 555	2,68	0,46
Québec	140 232	21 033	1,90	0,28
Ontario	436 349	33 472	3,68	0,27
Manitoba	30 903	5 423	2,69	0,46
Saskatchewan	21 231	4 333	2,12	0,42
Alberta	69 857	11 308	2,29	0,36
Colombie-Britannique	164 542	15 598	4,04	0,37
Yukon	1 423	372	4,73	1,18
Territoires du Nord-Ouest	3 295	362	8,11	0,82
Nunavut	1 256	411	4,49	1,40
Régions métropolitaines de recensement	626 971	38 179	3,15	0,19
St. John's	4 541	1 232	2,56	0,68
Halifax	8 097	2 736	2,20	0,73
Saint John	5 522	1 455	4,31	1,09
Chicoutimi-Jonquière ¹	610	2 610	0,39	1,67

Caractéristiques	Nombre net de personnes omises		Taux de sous-dénombrement net de la population	
	Nombre estimé (%)	Erreur-type	Taux estimé (%)	Erreur-type
Québec	14 598	6 930	2,09	0,97
Sherbrooke	580	2 560	0,38	1,65
Trois-Rivières	443	2 113	0,32	1,53
Montréal	67 291	15 150	1,93	0,43
Ottawa–Hull ²	44 633	11 258	4,03	0,97
Kingston	-3 993	2 920	-2,79	2,10
London	14 555	6 189	3,26	1,29
Oshawa	2 452	2 897	0,82	0,96
Toronto	255 368	24 343	5,17	0,47
Hamilton	11 519	5 561	1,71	0,81
St. Catharines–Niagara	24 042	6 526	5,99	1,53
Kitchener	9 746	5 581	2,30	1,29
Windsor	2 180	4 159	0,70	1,33
Greater Sudbury	7 802	4 712	4,77	2,75
Thunder Bay	1 819	1 965	1,47	1,56
Winnipeg	10 830	4 015	1,59	0,58
Regina	2 426	1 503	1,24	0,76
Saskatoon	2 768	2 255	1,21	0,97
Calgary	20 044	6 974	2,06	0,70
Edmonton	18 072	6 283	1,89	0,64
Abbotsford	8 481	2 858	5,44	1,73
Vancouver	79 757	11 884	3,86	0,55
Victoria	12 784	3 635	3,94	1,08
Âge et sexe				
Les deux sexes	924 429	44 749	2,99	0,14
0 à 4 ans	60 665	13 508	3,45	0,74
5 à 14 ans	56 502	16 992	1,38	0,41
15 à 19 ans	52 945	12 913	2,51	0,60
20 à 24 ans	151 443	16 120	7,19	0,71
25 à 34 ans	307 608	17 288	7,15	0,37
35 à 44 ans	187 782	18 640	3,55	0,34
45 à 54 ans	63 011	13 185	1,41	0,29
55 à 64 ans	24 280	11 044	0,84	0,38
65 ans et plus	20 194	15 833	0,52	0,40
Hommes	609 231	32 634	3,98	0,20

Caractéristiques	Nombre net de personnes omises		Taux de sous-dénombrement net de la population	
	Nombre estimé (%)	Erreur-type	Taux estimé (%)	Erreur-type
0 à 4 ans	23 854	8 201	2,67	0,89
5 à 14 ans	16 268	11 306	0,78	0,54
15 à 19 ans	44 224	9 877	4,03	0,86
20 à 24 ans	100 013	12 308	9,24	1,03
25 à 34 ans	209 784	13 773	9,64	0,57
35 à 44 ans	139 629	14 397	5,26	0,51
45 à 54 ans	48 187	10 029	2,16	0,44
55 à 64 ans	14 453	7 837	1,01	0,54
65 ans et plus	12 820	9 726	0,77	0,58
Femmes	315 199	31 393	2,02	0,20
0 à 4 ans	36 811	10 889	4,26	1,21
5 à 14 ans	40 234	12 730	2,01	0,62
15 à 19 ans	8 722	8 348	0,86	0,82
20 à 24 ans	51 430	10 484	5,02	0,97
25 à 34 ans	97 823	10 574	4,60	0,47
35 à 44 ans	48 153	11 903	1,83	0,44
45 à 54 ans	14 824	8 559	0,66	0,38
55 à 64 ans	9 826	7 775	0,67	0,53
65 ans et plus	7 375	12 487	0,33	0,56
État matrimonial légal et sexe				
Les deux sexes	924 429	44 750	2,99	0,14
Marié(e) ou séparé(e)	176 984	24 727	1,37	0,19
Divorcé(e)	64 975	10 811	3,38	0,54
Veuf(ve)	30 792	9 361	1,96	0,58
Jamais marié(e)	635 451	35 162	4,38	0,23
Indéterminé ⁴	16 227	3 503	•	•
Hommes	609 231	32 634	3,98	0,20
Marié ou séparé	117 052	17 816	1,81	0,27
Divorcé	56 126	8 567	6,49	0,93
Veuf	5 048	3 589	1,75	1,22
Jamais marié	420 671	26 023	5,46	0,32
Indéterminé ⁴	10 334	2 697	•	•
Femmes	315 199	31 393	2,02	0,20

Caractéristiques	Nombre net de personnes omises		Taux de sous-dénombrement net de la population	
	Nombre estimé (%)	Erreur-type	Taux estimé (%)	Erreur-type
Mariée ou séparée	59 932	17 229	0,93	0,26
Divorcée	8 849	6 606	0,84	0,62
Veuve	25 745	8 649	2,01	0,66
Jamais mariée	214 781	24 086	3,16	0,34
Indéterminé ⁴	5 893	2 254	.	.
État matrimonial et sexe				
Les deux sexes	924 429	44 750	2,99	0,14
Marié(e), séparé(e), ou union libre	236 196	26 864	1,56	0,17
Divorcé(e)	63 669	10 175	4,67	0,71
Veuf(ve)	31 725	9 352	2,10	0,61
Jamais marié(e)	576 613	33 832	4,47	0,25
Indéterminé ⁴	16 227	3 503	.	.
Hommes	609 231	32 634	3,98	0,20
Marié, séparé, ou union libre	154 771	19 500	2,04	0,25
Divorcé	51 996	8 037	8,96	1,26
Veuf	5 518	3 589	2,08	1,33
Jamais marié	386 613	25 032	5,61	0,34
Indéterminé ⁴	10 334	2 697	.	.
Femmes	315 199	31 393	2,02	0,20
Mariée, séparée, ou union libre	81 424	18 597	1,08	0,24
Divorcée	11 672	6 254	1,49	0,79
Veuve	26 207	8 640	2,11	0,68
Jamais mariée	190 002	23 165	3,16	0,37
Indéterminé ⁴	5 893	2 254	.	.
Union libre				
Les deux sexes	924 429	44 750	2,99	0,14
Vivant en union libre	62 481	11 132	2,66	0,46
Indéterminée ⁴	15 264	3 412	.	.
Hommes	609 231	32 634	3,98	0,20
Vivant en union libre	39 172	8 220	3,31	0,67
Indéterminée ⁴	9 371	2 578	.	.

Caractéristiques	Nombre net de personnes omises		Taux de sous-dénombrement net de la population	
	Nombre estimé (%)	Erreur-type	Taux estimé (%)	Erreur-type
Femmes	315 199	31 393	2,02	0,20
Vivant en union libre	23 309	7 496	2,01	0,63
Indéterminée ⁴	5 893	2 254	•	•
Langue maternelle				
Total	924 429	44 750	2,99	0,14
Français	124 431	19 837	1,84	0,29
Anglais	429 538	32 162	2,36	0,17
Autre que le français ou l'anglais	361 178	24 043	7,70	0,47
Français et anglais	3 024	4 121	0,97	1,31
Français et autre	-3 549	186	-3,60 ³	0,20
Anglais et autre	5 463	5 936	0,71	0,77
Français, anglais, et autre	-1 984	100	-2,69 ³	0,14
Indéterminée ⁴	6 328	3 481	•	•

• indisponible

1. Maintenant connue sous le nom de Saguenay.

2. Maintenant connue sous le nom d'Ottawa–Gatineau.

3. Les chiffres et taux estimés négatifs révèlent un surdénombrement net.

4. Les taux estimés et erreurs-types correspondantes ne sont pas établis pour la catégorie « Indéterminé » ou « Indéterminée ».

5.1 Sous-dénombrement net de la population

Le sous-dénombrement net de la population est plus élevé dans l'Ouest.

Les provinces affichant les taux de sous-dénombrement net de la population les plus élevés sont la Colombie-Britannique (4,04 %), suivie de l'Ontario (3,68 %) et du Manitoba (2,69 %). Dans l'ensemble du pays, ce sont les Territoires du Nord-Ouest qui présentent le taux le plus élevé (8,11 %). L'Île-du-Prince-Édouard (0,97 %), Terre-Neuve-et-Labrador (1,80 %) et le Québec (1,90 %) présentent les taux les plus faibles.

Le sous-dénombrement net de la population est le plus élevé chez les jeunes adultes âgés de 20 à 34 ans.

Le taux de sous-dénombrement net élevé chez les jeunes adultes âgés de 20 à 34 ans témoigne de la forte mobilité de résidence des personnes de ce groupe d'âge. Il s'établit à 7,19 % pour les personnes

âgées de 20 à 24 ans et ne baisse que légèrement pour les personnes âgées de 25 à 34 ans (7,15 %). Les taux de sous-dénombrement net de ces groupes d'âge dépassent de beaucoup le taux national tant chez les hommes que chez les femmes. On observe les taux de sous-dénombrement net les plus élevés chez les hommes âgés de 25 à 34 ans et de 20 à 24 ans, soit 9,64 % et 9,24 % respectivement. Pour tous les groupes d'âge combinés, le sous-dénombrement net est plus marqué chez les hommes (3,98 %) que chez les femmes (2,02 %). De fait, les taux des hommes dépassent ceux des femmes pour tous les groupes d'âge supérieur à 14 ans.

Le sous-dénombrement net est plus élevé chez les personnes divorcées et les personnes jamais mariées.

La différence globale de 1,96 point de pourcentage entre le taux des hommes et celui des femmes s'explique en grande partie par les taux des personnes jamais mariées et des personnes divorcées. En effet, le taux de sous-dénombrement net des hommes divorcés atteint 6,49 % comparativement à 0,84 % pour les femmes divorcées. Dans le cas des hommes jamais mariés, le taux s'établit à 5,46 % contre 3,16 % pour les femmes du même groupe. Lorsqu'on utilise la variable « état matrimonial », qui, contrairement à la variable « état matrimonial légal », inclut les couples en union libre dans la catégorie des couples mariés, on observe les taux de sous-dénombrement net les plus importants chez les personnes divorcées (4,67 %) et chez les personnes jamais mariées (4,47 %).

Le sous-dénombrement net des personnes vivant en union libre est légèrement inférieur à celui de l'ensemble de la population.

Le taux de sous-dénombrement net des hommes vivant en union libre se fixe à 3,31 % comparativement à 3,98 % pour l'ensemble des hommes. Dans le cas des femmes vivant en union libre, le taux de sous-dénombrement net atteint 2,01 % contre 2,02 % pour l'ensemble des femmes.

Les personnes dont la langue maternelle est « autre que le français ou l'anglais » affichent le taux de sous-dénombrement net le plus élevé.

On observe des différences entre le taux de sous-dénombrement net des personnes ayant indiqué l'anglais comme langue maternelle (2,36 %) et celui des personnes ayant indiqué le français comme langue maternelle (1,84 %). Toutefois, le taux augmente considérablement dans le cas de personnes ayant indiqué une langue non officielle comme langue maternelle (7,70 %).

6. Programme de mesure de l'erreur de couverture

6.1 Portée et objectifs

Le Programme de mesure de l'erreur de couverture de 2001 porte principalement sur l'univers de la population et l'univers des logements. Il permet de mesurer les composantes suivantes de l'erreur de couverture :

- les erreurs de classement relatives aux logements privés inoccupés et aux logements non répondants du recensement;
- le sous-dénombrement et le surdénombrement de la population.

Le Programme de mesure de l'erreur de couverture de 2001 comportait quatre études :

- l'Étude sur la classification des logements (ÉCL);
- la Contre-vérification des dossiers (CVD);
- l'Étude par appariement automatisé (ÉAA);
- l'Étude sur les logements collectifs (ÉLC).

L'ÉCL a permis de produire des estimations du sous-dénombrement découlant du classement erroné de logements dans la catégorie des logements inoccupés ainsi que des non répondants au recensement. La CVD visait à estimer le sous-dénombrement total provenant de toutes les sources, y compris le sous-dénombrement mesuré par l'ÉCL. Par ailleurs, la CVD, l'ÉAA et l'ÉLC ont permis de mesurer le surdénombrement. L'ÉAA consistait à repérer les personnes dénombrees dans deux ménages ou plus dans une même région (Atlantique, Québec, Ontario, et Ouest canadien comprenant la Colombie-Britannique et les trois territoires). L'ÉLC visait à estimer le surdénombrement découlant du fait que des personnes dénombrees comme résidents habituels dans un logement collectif non institutionnel étaient également dénombrees dans un logement privé. La CVD avait pour but d'évaluer le surdénombrement provenant de toutes les sources. Si l'ÉAA ou l'ÉLC révèle un surdénombrement, celui-ci est enlevé de la CVD, de façon à éviter que le surdénombrement ne soit compté plus d'une fois par les trois études.

Tableau 6.1 Programme de mesure de l'erreur de couverture au recensement de 2001

Étude	Taille de l'échantillon	Objectif
Étude sur la classification des logements	1 399 secteurs de dénombrement	Permet de mesurer le sous-dénombrement découlant des erreurs de classification de logements comme logements inoccupés et ainsi que des <i>logements</i> non répondants au recensement.
Contre-vérification des dossiers	60 653 personnes	Permet de mesurer le sous-dénombrement provenant de toutes les sources et le surdénombrement non mesuré par l'Étude par appariement automatisé ni par l'Étude sur les logements collectifs.
Étude par appariement automatisé	17 275 paires de ménages	Permet de mesurer le surdénombrement découlant du fait que des personnes ont été dénombrees dans deux ménages ou plus résidant dans la même région.
Étude sur les	4 500 résidents	Permet de mesurer le surdénombrement imputable au

Étude	Taille de l'échantillon	Objectif
logements collectifs	habituels	fait que des personnes ont été dénombrées dans un logement collectif non institutionnel et dans un logement privé.

Les données obtenues à l'aide de ces quatre études ont été utilisées comme suit :

- Les estimations obtenues à l'aide de l'ÉCL ont été incluses dans les chiffres définitifs du recensement pour tenir compte de cette source précise du sous-dénombrement.
- Les estimations obtenues dans le cadre de la CVD, de l'ÉAA et de l'ÉLC n'ont pas été incluses dans les chiffres du recensement. Elles sont incluses dans la population de base pour le Programme des estimations de la population de Statistique Canada.
- Les renseignements sur les causes et les caractéristiques des erreurs de couverture qui permettent de repérer les régions ou les sous-groupes de population affichant des taux particulièrement élevés d'erreur de couverture servent à la planification du prochain recensement.
- Les renseignements supplémentaires recueillis par les études ont servi à l'évaluation de la qualité de certaines questions du recensement. Ainsi par exemple, la CVD permet de coupler les réponses d'une personne à deux recensements successifs de sorte qu'il est possible de calculer l'erreur de réponse à certaines questions comme celles portant sur la date de naissance, le sexe et la langue maternelle.
- Enfin, les résultats servent à informer les utilisateurs sur la nature et l'ampleur des erreurs de couverture du recensement de sorte qu'ils puissent tirer des conclusions à partir des données du recensement ou prendre des décisions fondées sur celles-ci en toute connaissance de cause.

La méthodologie et les résultats de chacune des études de mesure de la couverture sont présentés dans les chapitres suivants.

6.2 Améliorations

Les changements suivants ont été apportés aux études de la couverture de 1996 afin d'en améliorer la qualité et d'obtenir une mesure plus exacte de l'erreur de couverture de 2001.

- La Vérification des logements inoccupés (VLI) a été remplacée par l'Étude sur la classification des logements (ÉCL). Cette nouvelle étude est utilisée pour effectuer la correction des chiffres du recensement avant publication, pour les logements classés à tort comme logements inoccupés et les logements identifiés comme étant occupés pour lesquels le recensement n'a pas réussi à établir un contact, en recueillant de l'information sur les caractéristiques des occupants de ces logements (aussi appelés logements non répondants). L'ÉCL est plus exhaustive et plus précise que l'ancienne VLI.
- La collecte des données aux fins de la Contre-vérification des dossiers (CVD) a été effectuée grâce au procédé de l'interview téléphonique assistée par ordinateur (ITAO). Contrairement au questionnaire papier, qui était utilisé par le passé, l'ITAO offre la possibilité d'effectuer des vérifications automatisées ainsi que la correction immédiate des données. Ainsi, le recours à l'ITAO a permis d'améliorer la qualité des données. Il a également permis de réduire de façon considérable les

délais de transfert des données au bureau central et de soumission des données à une série de vérifications supplémentaire.

- Dans le cas de la Contre-vérification des dossiers (CVD), on a automatisé le système de traitement des adresses des personnes choisies. Ce système sert à déterminer si les personnes faisant partie de l'échantillon ont été dénombrées ou non aux adresses recueillies. Pour effectuer le traitement, les commis avaient accès à toute l'information nécessaire depuis leur poste de travail (p. ex., informations concernant les personnes et les adresses, bottins téléphoniques électroniques, cartes géographiques). Grâce à l'efficacité accrue de ce système, on a pu traiter un plus grand nombre d'adresses et réduire de façon significative le nombre d'adresses qui n'étaient toujours pas résolues à la fin du traitement. Ces résultats favorables sont aussi en partie attribuables à l'utilisation d'outils de recherche plus sophistiqués comme des bottins téléphoniques et des cartes géographiques électroniques.
- La CVD de 2001 a permis de produire des estimations pour toutes les erreurs de couverture, y compris celles qui avaient déjà été corrigées dans les chiffres publiés du recensement. Les ajustements aux chiffres du recensement avant publication (faits en fonction des résultats de l'ÉCL) ont été supprimés et les résultats de la CVD utilisés pour tous les cas de sous-dénombrement.
- On a abandonné la composante institutionnelle de l'Étude sur les logements collectifs (une étude sur le surdénombrement) et produit les estimations du surdénombrement chez ce groupe de personnes (très petit historiquement) dans le cadre de la CVD.

7. Étude sur la classification des logements

7.1 Introduction

Le classement erroné de logements constitue l'une des sources possibles d'erreurs dans le recensement. Une telle erreur se produit si un logement occupé est classé comme étant « inoccupé », ou si un logement inoccupé est classé comme étant « occupé ». Si un logement occupé est classé par erreur comme étant inoccupé, il s'ensuit un sous-dénombrement des personnes et des logements occupés dans les chiffres du recensement. En revanche, si un logement inoccupé est classé comme étant occupé, il se produit un surdénombrement des personnes et des logements occupés. Ce dernier genre d'erreur de classement peut survenir si le logement a été classé par le recenseur comme étant occupé alors qu'aucun questionnaire n'a été retourné. Les logements de ce type sont appelés logements non répondants.

On crée aussi un surdénombrement, du parc immobilier cette fois, si l'on inclut à tort les logements marginaux ou les logements en construction dans la classe des logements inoccupés.

L'Étude sur la classification des logements (ÉCL) vise à étudier ces types d'erreurs de classement. Il s'agit d'une version élargie de la Vérification des logements inoccupés qui servait, lors des recensements antérieurs, à réexaminer les logements que les recenseurs avaient classés comme étant inoccupés.

Les renseignements recueillis à l'aide de l'ÉCL de 2001 servent à :

- évaluer le nombre de logements inoccupés étant exclus de l'univers des logements; il s'agit des logements marginaux, des logements en construction et des immeubles qui ne sont pas des logements (stations-services, cabinets de médecin, etc.);
- évaluer le nombre de logements occupés qui ont été classés à tort comme étant inoccupés pendant le recensement et à évaluer le nombre de ménages et de personnes omis en raison de cette erreur de classement;
- **ajuster les données du recensement sur les ménages et les personnes afin de corriger cette erreur de classement;**
- estimer le nombre de logements non répondants qui étaient inoccupés;
- estimer le nombre de personnes vivant dans les logements non répondants;
- **ajuster la répartition de la taille des ménages en procédant par imputation pour les logements non répondants.**

Les trois derniers points sont des éléments nouveaux pour l'ÉCL de 2001, alors que la portée de l'étude a été élargie afin d'inclure l'examen des logements non répondants en plus des logements inoccupés.

7.2 Méthodologie

7.2.1 Stratification et sélection de l'échantillon

La population visée par l'Étude sur la classification des logements (ÉCL) comprend l'ensemble des logements non répondants et des logements inoccupés repérés au recensement du 15 mai 2001, sauf les logements dans les secteurs de dénombrement (SD) collectifs, dans les SD dénombrés par interview et dans les réserves indiennes. Ces derniers ont été exclus de la base de sondage, principalement en raison de considérations d'ordre opérationnel et pécuniaire.

On a établi la taille de l'échantillon visé par l'ÉCL de 2001 à 1 399 SD répartis dans l'ensemble du Canada. La base de sondage comprenait tous les SD de retour par la poste, sauf les réserves indiennes.

Par conséquent, les régions rurales du Yukon et des Territoires du Nord-Ouest et l'ensemble du Nunavut sont exclus de la base de sondage. L'échantillon établi pour les territoires ne comprend que Whitehorse, Yellowknife, Hay River et Fort Smith.

La sélection préliminaire de l'échantillon a eu lieu pour choisir des SD qui avaient uniquement fait l'objet de la méthode de collecte de retour par la poste et qui n'étaient pas liés à des réserves indiennes, des établissements indiens ou d'autres types de subdivisions de recensement indiennes. Les SD étaient ensuite répartis entre la base de sondage des SD urbains et des SD ruraux. Pour être inclus dans la base de sondage des SD urbains, un SD devait d'abord faire partie d'une agglomération de recensement (AR) ou d'une région métropolitaine de recensement (RMR) comprenant 50 000 logements occupés ou plus. Si plus de 50 % des SD situés dans un district de commissaire au recensement (DCR) au sein d'une AR/RMR étaient classés comme urbains, alors la totalité des SD dans le DCR étaient classés dans la base de sondage des SD urbains, et l'échantillon était tiré selon une méthode d'échantillonnage à un degré. Autrement, tous les SD dans le DCR étaient considérés comme des SD ruraux, et on utilisait une méthode d'échantillonnage à deux degrés. Tous les SD ne faisant pas partie de la base de sondage des SD urbains étaient inclus dans la base de sondage des SD ruraux. L'échantillon visé par l'ÉCL a alors été constitué à partir de ces deux bases.

Échantillon urbain (un degré)

Dans les régions compactes, qui tendaient principalement à se trouver dans les centres urbains, on utilisait une méthode d'échantillonnage à un degré. L'échantillon urbain comportait trois composantes distinctes. Dans le Yukon (Whitehorse seulement) et les Territoires du Nord-Ouest (Yellowknife, Hay River et Fort Smith seulement), tous les SD dans la base de sondage ont été sélectionnés pour l'ÉCL. Pour l'Île-du-Prince-Édouard, un échantillon aléatoire simple de 45 SD a été tiré. Pour sélectionner l'échantillon de SD urbains dans toutes les autres provinces, on a stratifié les SD urbains selon l'AR ou la RMR au sein de chaque province, c'est-à-dire que chaque AR/RMR constituait une strate, et un échantillon aléatoire simple du nombre requis de SD a été tiré dans chacune des strates. Dans chaque strate, chaque échantillon comptait un minimum de cinq SD. On a ainsi obtenu 678 SD urbains en tout dans l'échantillon.

Échantillon rural (deux degrés)

La sélection de l'échantillon rural s'est faite à l'aide d'une méthode d'échantillonnage à deux degrés, et ce, afin de réduire les coûts sur le terrain. Les coûts de l'ÉCL afférents aux interviews sur le terrain, particulièrement les frais de déplacement, peuvent augmenter considérablement en dehors des régions urbaines. À partir des données de recensements antérieurs (1986, 1991 et 1996), on estime que cinq SD regroupés constituent une tâche d'ampleur raisonnable pour un intervieweur affecté à l'ÉCL. Afin de regrouper cinq SD assez près les uns des autres de manière à former une unité relativement compacte, on a effectué l'échantillonnage à deux degrés. Dans un premier temps, le nombre alloué de DCR a été sélectionné de façon aléatoire dans chaque province (secteurs ruraux). Dans un deuxième temps, cinq SD ont été choisis de façon aléatoire dans chaque DCR choisi. Cette méthode d'échantillonnage a permis d'obtenir 721 SD dans l'échantillon rural.

L'échantillon de l'ÉCL comprenait tous les logements inoccupés ainsi que tous les logements non répondants dans les SD échantillonnés. Ainsi, au total, 29 777 logements inoccupés et 2 749 logements non répondants ont été sélectionnés pour faire partie de l'échantillon de l'ÉCL. Le tableau 7.2.1 montre la répartition de l'échantillon selon la province ou le territoire.

Tableau 7.2.1 Taille de l'échantillon, Canada, provinces et territoires, Étude sur la classification des logements de 2001

	N ^{bre} de secteurs de dénombrement dans l'échantillon	N ^{bre} de logements inoccupés dans l'échantillon	N ^{bre} de logements non répondants dans l'échantillon
Canada	1 399	29 777	2 749
Terre-Neuve-et-Labrador	80	2 635	58
Île-du-Prince-Édouard	45	874	1
Nouvelle-Écosse	96	2 501	287
Nouveau-Brunswick	65	1 090	124
Québec	271	6 700	860
Ontario	278	4 513	444
Manitoba	94	2 991	90
Saskatchewan	126	2 400	91
Alberta	125	2 157	276
Colombie-Britannique	161	2 733	455
Yukon	28	330	55
Territoires du Nord-Ouest	30	853	8
Nunavut	0	0	0

7.2.2 Interviews sur le terrain

Chaque logement dans un SD échantillonné classé soit comme étant inoccupé le jour du recensement, soit comme étant occupé mais pour lequel aucun questionnaire n'a été retourné a été vérifié de nouveau à la fin juin ou au début juillet 2001, en vue de déterminer s'il était réellement inoccupé/occupé le jour du recensement. Un questionnaire de l'Étude sur la classification des logements a été conçu précisément à cette fin.

On a laissé chaque bureau régional choisir quel était le moment propice pour mener cette opération après le recensement. Pour déterminer l'occupation/inoccupation et dans le but d'obtenir d'autres renseignements, on demandait aux recenseurs de communiquer avec les occupants actuels ou avec un voisin, le propriétaire ou une autre personne connaissant le logement en question. Pour chaque logement, le recenseur pouvait faire jusqu'à trois tentatives de contacts. S'il constatait que le logement avait été occupé le jour du recensement, il obtenait également le nombre d'occupants au jour du recensement et leurs noms.

7.2.3 Dépouillement, codage et contrôle

Une fois les interviews sur le terrain terminées, tous les questionnaires ont été envoyés à Ottawa en vue de leur dépouillement. Une fois rendus au bureau central, les questionnaires ont fait l'objet d'un dépouillement préliminaire, puis on en a ensuite fait la saisie sur clavier.

D'abord, les questionnaires ne faisant pas partie de l'échantillon étaient éliminés, et lorsque plus d'un questionnaire provenait d'une même adresse, on obtenait le questionnaire approprié pour chaque logement. On faisait ensuite des contrôles préliminaires et le pré-traitement général des questionnaires avant de les envoyer aux opérations de saisie de données.

Une fois la saisie des données terminée, les questionnaires ont fait l'objet d'un ensemble complet de contrôles de la cohérence. On a examiné à la main les questionnaires rejetés au contrôle afin de résoudre les incohérences.

Dans le cas de chaque logement pour lequel on déterminait qu'il avait été occupé le jour du recensement, on vérifiait le questionnaire de l'Étude de la classification des logements (ÉCL) afin de voir s'il renfermait une autre adresse où les membres du ménage pouvaient avoir été dénombrés. Si une autre adresse était fournie, on vérifiait le Registre des visites et d'autres questionnaires du recensement afin de déterminer si les membres du ménage avaient effectivement été dénombrés ailleurs. Le cas échéant, ils n'étaient pas ajoutés au chiffre des personnes, mais le logement lui-même était ajouté au chiffre des logements occupés.

À partir de là, les logements inoccupés et les logements non répondants n'étaient plus traités de la même façon. La liste des questionnaires de l'ÉCL remplis pour chaque secteur de dénombrement échantillonné était appariée à la liste finale des logements inoccupés selon le recensement, de façon à créer le fichier final des logements inoccupés devant être traité par la suite. Les logements pour lesquels un questionnaire de l'ÉCL avait été reçu mais qui n'étaient pas inscrits au recensement étaient retirés du champ de l'étude. Les logements qui étaient inscrits au recensement comme étant inoccupés mais pour lesquels on n'avait reçu aucun questionnaire de l'ÉCL étaient considérés comme des logements non répondants aux fins de cette étude. De même, la liste des questionnaires de l'ÉCL était appariée à la liste finale de tous les logements pour lesquels on n'avait reçu aucun questionnaire du recensement, de façon à créer le fichier final des logements non répondants devant être traité par la suite.

7.2.4 Non-réponse, imputation et pondération

Bien que les logements inoccupés et les logements non répondants soient traités séparément, les processus de traitement de la non-réponse, d'imputation et de pondération étaient identiques dans les deux cas.

La non-réponse totale (c.-à-d. aucun renseignement obtenu sur un logement donné) a été traitée par ajustement des poids au sein de plusieurs régions infraprovinciales, soit les trois plus grandes régions métropolitaines de recensement (Montréal, Toronto et Vancouver), ainsi que les autres parties urbaines et rurales de chaque province et territoire.

Pour régler les cas de non-réponse partielle sur l'occupation/inoccupation, le nombre de résidents habituels et le type de logement, on a eu recours à l'imputation. On a imputé en premier lieu l'occupation/inoccupation, dont on s'est servi ensuite pour imputer les autres variables.

Les poids ont ensuite été ajustés de sorte que leur somme égale le nombre connu de logements inoccupés ou non répondants repérés lors du recensement dans chacune des régions infraprovinciales.

La dernière étape du traitement de l'Étude sur la classification des logements (ÉCL) consistait à redresser les bases de données du recensement. Dans le cas des logements inoccupés, il a d'abord fallu préparer un profil à l'échelle nationale des logements incorrectement classés pour les régions urbaines et rurales, en se servant du type de logement et du nombre de personnes omises en raison de l'erreur de classification. Ces profils nationaux ont ensuite servi au calcul d'estimations du nombre de logements classés incorrectement selon le nombre de personnes dans le logement, le type de logement et les régions rurales ou urbaines à l'échelle des provinces et des territoires. En se fondant sur ces estimations, on a sélectionné au hasard des logements dénombrés présentant les mêmes caractéristiques (nombre de personnes, type de logement privé) et augmenté d'une unité leur poids dans le recensement. Pour chaque ménage choisi, le poids d'un logement inoccupé dans le même SD a été réduit à zéro de sorte que le nombre total de logements n'augmente pas.

Les chiffres du recensement ont été ajustés de la façon suivante en ce qui touche le nombre de personnes vivant dans les logements non répondants. Si le recenseur avait inscrit la taille du ménage sur

le questionnaire du recensement, celle-ci était estimée exacte. Dans le cadre de l'ÉCL, on établissait le profil des logements non répondants, qui fournissait le nombre estimé de logements pour chaque catégorie de taille du ménage (de zéro à six personnes) pour chaque région géographique. Une seconde distribution était établie à partir des données observées du recensement, puis un vecteur de classement erroné (différence entre les deux distributions) était créé. Des ménages non répondants étaient alors choisis au hasard parmi ceux disponibles dans chaque région géographique et une nouvelle taille leur était imputée jusqu'à ce que le vecteur de classement erroné soit égal à zéro (c'est-à-dire jusqu'à ce que la distribution des tailles des ménages non répondants au recensement soit égale à la distribution des tailles des ménages non répondants établie selon l'ÉCL).

7.3 Résultats

Les principaux résultats ayant trait aux logements classés initialement comme étant inoccupés au recensement de 2001 sont présentés dans les tableaux 7.3.1A, 7.3.1C, 7.3.1D et 7.3.1E. Le tableau 7.3.1A fournit le nombre estimé et le taux de logements classés comme étant inoccupés qui, de fait, étaient occupés, selon qu'ils sont situés dans une région urbaine ou rurale, selon la région, la province, la région métropolitaine de recensement choisie et le type de logement privé. Le tableau 7.3.1B donne les mêmes renseignements, mais selon la Vérification des logements inoccupés de 1996. Le tableau 7.3.1C présente le nombre de ménages et de personnes ajoutés aux chiffres du recensement de 2001 à cause de ces erreurs de classement et le tableau 7.3.1D, le nombre de logements inoccupés hors du parc de logements selon les mêmes répartitions que celles du tableau 7.3.1A. Enfin, le tableau 7.3.1E fait état de l'erreur de classement des logements au recensement qui est attribuable au fait que des logements qui n'auraient pas dû être inclus dans le parc de logements ont été classés par erreur comme étant inoccupés.

Les résultats obtenus pour les logements non répondants (logements considérés comme étant occupés mais pour lesquels on n'a reçu aucun questionnaire du recensement) sont présentés au tableau 7.3.2, qui fait état du nombre de personnes vivant dans ces logements (estimé selon l'ÉCL), ainsi que de l'erreur-type estimée.

7.3.1 Logements inscrits comme étant inoccupés au recensement

A. Logements occupés

Le tableau 7.3.1A montre qu'à l'échelle du pays 12,7 % des logements classés comme étant inoccupés lors du recensement étaient en fait occupés. Il s'agit aussi de logements occupés par des résidents étrangers ou temporaires uniquement ainsi que de logements pour lesquels au moins une personne avait été dénombrée ailleurs au Canada. Ce pourcentage constitue une hausse par rapport à celui de 7,8 % enregistré en 1996 (voir le tableau 7.3.1B). Le classement erroné de logements est beaucoup plus fréquent dans les régions urbaines (21,3 %) que dans les régions rurales (5,7 %). Le taux calculé pour les régions rurales diffère très peu du taux affiché en 1996, alors que celui des régions urbaines présente une forte hausse par rapport au taux de 9,9 % enregistré en 1996. Cette augmentation observée dans les régions urbaines entre 1996 et 2001 est attribuable en grande partie à des hausses survenues dans les trois plus grandes régions métropolitaines de recensement (RMR) (Montréal, Toronto et Vancouver), surtout dans les grands immeubles d'appartements.

Tableau 7.3.1A Nombre estimé de logements occupés classés par erreur comme étant inoccupés au recensement de 2001, Étude sur la classification des logements de 2001¹

Région / Type de logement privé	N ^{bre} de logements classés comme logements inoccupés au recensement de 2001	Logements occupés			
		Total estimé	Erreur-type	Taux (%)	Erreur-type (%)
Canada	904 236	114 603	7 166	12,7	0,8
Urbain	405 427	86 370	5 814	21,3	1,4
Rural	498 809	28 233	4 176	5,7	0,8
Provinces de l'Atlantique	112 357	6 961	693	6,2	0,6
Terre-Neuve-et-Labrador	35 896	1 511	291	4,2	0,8
Île-du-Prince-Édouard	5 121	396	100	7,7	2,0
Nouvelle-Écosse	42 931	3 411	519	7,9	1,2
Nouveau-Brunswick	28 409	1 643	341	5,8	1,2
Québec	213 062	31 007	4 577	14,6	2,1
Ontario	316 474	36 647	4 676	11,6	1,5
Prairies	150 294	15 303	1 729	10,2	1,2
Manitoba	37 592	2 737	434	7,3	1,2
Saskatchewan	47 503	3 113	486	6,6	1,0
Alberta	65 199	9 452	1 602	14,5	2,5
Colombie-Britannique	110 755	24 545	2 252	22,2	2,0
Territoires	1 294	140	27	10,8	2,1
Yukon	395	72	10	18,2	2,5
Territoires du Nord-Ouest	899	68	25	7,6	2,8
Certaines RMR					
Montréal	42 174	8 435	1 737	20,0	4,1
Toronto	40 020	17 557	4 330	43,9	10,8
Vancouver	29 565	11 034	1 614	37,3	5,5
Type de logement privé					
Maison individuelle non attenante	423 328	46 616	3 554	11,0	0,8
Appartement dans un immeuble de cinq étages ou plus	39 365	18 796	4 427	47,7	11,2
Autre type de logement	230 899	49 191	4 434	21,3	1,9
Hors du parc de logements	210 644

... n'ayant pas lieu de figurer

(1) Ne s'applique pas au Nunavut.

Tableau 7.3.1B Nombre estimé de logements occupés classés par erreur comme étant inoccupés au recensement de 1996, Vérification des logements inoccupés de 1996

Région / Type de logement privé	N ^{bre} de logements classés comme logements inoccupés au recensement de 1996	Logements occupés			
		Total estimé	Erreur-type	Taux (%)	Erreur-type (%)
Canada	781 594	61 287	3 296	7,8	0,4
Urbain	415 474	41 295	2 570	9,9	0,6
Rural	366 120	19 992	2 021	5,5	0,6
Provinces de l'Atlantique	67 671	3 303	488	4,9	0,7
Terre-Neuve	23 081	1 001	303	4,3	1,3
Île-du-Prince-Édouard	3 109	176	40	5,7	1,3
Nouvelle-Écosse	25 842	1 305	291	5,1	1,1
Nouveau-Brunswick	15 639	822	246	5,3	1,6
Québec	216 838	13 298	1 502	6,1	0,7
Ontario	262 721	26 611	2 457	10,1	0,9
Prairies	142 773	9 187	849	6,4	0,6
Manitoba	32 598	1 900	326	5,8	1,0
Saskatchewan	40 276	1 757	341	4,4	0,8
Alberta	69 899	5 530	706	7,9	1,0
Colombie-Britannique	90 112	8 714	1 269	9,7	1,4
Territoires	1 479	173	32	11,7	2,2
Yukon	748	147	30	19,7	4,0
Territoires du Nord-Ouest	731	26	10	3,6	1,4
Certaines régions métropolitaines de recensement					
Montréal	70 552	5 093	858	7,2	1,2
Toronto	38 210	8 259	1 588	21,6	4,2
Vancouver	29 711	3 737	765	12,6	2,6
Type de logement privé					
Maison individuelle non attenante	329 517	24 987	2 172	7,6	0,7
Appartement dans un immeuble de cinq étages ou plus	62 829	7 593	1 584	12,1	2,5
Autre type de logement	272 067	28 707	2 102	10,6	0,8

À l'échelle des provinces et territoires, la Colombie-Britannique montre le taux d'erreur de classement le plus élevé (22,2 %). Il était suivi du Yukon (18,2 %), du Québec (14,6 %), de l'Alberta (14,5 %) et de

l'Ontario (11,6 %). Le taux d'erreur de classement des autres provinces étaient assez comparables, allant de 4,2 % à Terre-Neuve-et-Labrador à 7,9 % en Nouvelle-Écosse.

Dans les trois plus grandes RMR, où le taux de classement erroné de logements de 2001 était très élevé, le taux à Toronto (43,9 %) était plus élevé qu'à Vancouver (37,3 %) ou Montréal (20,0 %).

Quant aux types de logement classés au recensement, le taux d'erreur de classement le plus bas touche les maisons individuelles non attenantes (11,0 %) et le taux le plus élevé touche les appartements dans un immeuble de cinq étages ou plus (47,7 %). Le taux d'erreur de classement est de 21,3 % dans la catégorie « autre type de logement », qui comprend les maisons jumelées, les maisons en rangée, les duplex, les appartements dans un immeuble de moins de cinq étages, les habitations mobiles et les autres logements mobiles.

Étant donné ces erreurs de classement, un certain nombre de ménages et de personnes n'ont pas été dénombrés au recensement de 2001. Toutefois, quelques-uns des 114 603 logements que l'on estime avoir été classés à tort comme étant inoccupés avaient en fait aussi été correctement dénombrés par le recenseur comme logements occupés (ils étaient inscrits deux fois dans le Registre des visites) et d'autres étaient occupés uniquement par des résidents étrangers ou temporaires qui ne devaient pas être inclus (à raison) dans les chiffres du recensement. Par conséquent, on a estimé que le nombre réel de ménages non dénombrés était de 111 628 et ce nombre de ménages a été ajouté aux chiffres du recensement par le biais de l'Étude sur la classification des logements (ÉCL). Le tableau 7.3.1C montre les nombres réels de ménages et de personnes ajoutés aux chiffres du recensement par l'ÉCL.

Tableau 7.3.1C Nombre de ménages et de personnes dans des logements classés par erreur comme étant inoccupés, qui ont été ajoutés aux chiffres du recensement de 2001 à la suite de l'Étude sur la classification des logements de 2001¹

Région / Type de logement privé	N ^{bre} de logements classés comme logements inoccupés au recensement de 2001	Ménages ajoutés		Personnes ajoutées	
		Total	Erreur-type	Total	Erreur-type
Canada	904 236	111 628	6 815	222 720	11 711
Urbain	405 427	84 110	5 418	170 410	9 144
Rural	498 809	27 518	4 124	52 310	7 298
Provinces de l'Atlantique	112 357	6 738	672	12 219	1 081
Terre-Neuve-et-Labrador					
Labrador	35 896	1 407	258	2 439	419
Île-du-Prince-Édouard	5 121	384	99	768	184
Nouvelle-Écosse	42 931	3 304	509	6 058	756
Nouveau-Brunswick	28 409	1 643	341	2 954	622
Québec	213 062	29 370	4 126	54 267	6 782
Ontario	316 474	36 188	4 617	78 393	8 204
Prairies	150 294	15 047	1 721	29 610	3 293
Manitoba	37 592	2 648	437	4 663	692
Saskatchewan	47 503	2 977	472	5 383	943
Alberta	65 199	9 422	1 597	19 564	3 078
Colombie-Britannique	110 755	24 145	2 164	47 933	3 440

Région / Type de logement privé	N ^{bre} de logements classés comme logements inoccupés au recensement de 2001	Ménages ajoutés		Personnes ajoutées	
		Total	Erreur-type	Total	Erreur-type
Territoires	1 294	140	27	298	55
Yukon	395	72	10	151	24
Territoires du Nord-Ouest	899	68	25	147	50
Certaines régions métropolitaines de recensement					
Montréal	42 174	8 392	1 720	16 231	3 142
Toronto	40 020	17 274	4 292	40 106	8 413
Vancouver	29 565	10 845	1 586	22 869	3 123
Type de logement privé					
Maison individuelle non attenante	423 328	45 855	3 525	106 434	7 801
Appartement dans un immeuble de cinq étages ou plus	39 365	17 315	3 894	28 028	6 011
Autre type de logement	230 899	48 458	4 367	88 257	8 024

(1) Ne s'applique pas au Nunavut.

B. Logements hors du parc de logements

Le dénombrement de logements inoccupés ne faisant pas partie de l'univers des logements entraîne le surdénombrement des logements. Les logements utilisés à des fins commerciales, ou qui ne sont pas habitables durant toute l'année, ou qui sont comptés deux fois lors du recensement sont considérés comme ne faisant pas partie de l'univers des logements. Un logement est compté deux fois s'il est inscrit à la fois comme étant occupé et inoccupé dans le Registre des visites.

Pour être considéré comme propre à l'habitation durant toute l'année, un logement doit être doté d'une source de chauffage ou d'énergie et fournir un espace clos permettant de s'abriter contre les intempéries. Il est parfois difficile de déterminer si un logement est habitable, par exemple, dans le cas de chalets, de logements en construction presque terminée ou de logements détériorés. Cette considération peut ainsi avoir une certaine part de subjectivité, de sorte que divers recenseurs peuvent classer différemment un même logement. Pour cette raison, les estimations du nombre de logements inoccupés pour lesquels on a déterminé lors de l'Étude sur la classification des logements qu'ils ne faisaient pas partie du parc de logements, estimations présentées dans le tableau 7.3.1D, devraient être utilisées avec prudence. La base de données du recensement n'est pas corrigée en fonction de ces valeurs.

Tableau 7.3.1D Estimation du nombre de logements inoccupés hors du parc de logements, Étude sur la classification des logements de 2001¹

Région	N ^{bre} de logements classés comme logements inoccupés au recensement de 2001	Logements hors du parc de logements			
		Total estimé	Erreur-type	Taux (%)	Erreur-type (%)
Canada	904 236	210 644	15 264	23,3	1,7
Urbain	405 427	82 915	7 368	20,5	1,8
Rural	498 809	127 729	13 239	25,6	2,7
Provinces de l'Atlantique	112 357	16 710	2 401	14,9	2,1
Terre-Neuve-et-Labrador	35 896	3 424	776	9,5	2,2
Île-du-Prince-Édouard	5 121	2 132	436	41,6	8,5
Nouvelle-Écosse	42 931	5 748	926	13,4	2,2
Nouveau-Brunswick	28 409	5 406	2 030	19,0	7,1
Québec	213 062	73 410	9 422	34,5	4,4
Ontario	316 474	52 008	8 298	16,4	2,6
Prairies	150 294	45 914	5 103	30,5	3,4
Manitoba	37 592	8 466	2 328	22,5	6,2
Saskatchewan	47 503	13 837	2 584	29,1	5,4
Alberta	65 199	23 611	3 734	36,2	5,7
Colombie-Britannique	110 755	22 108	6 598	20,0	6,0
Territoires	1 294	493	159	38,1	12,3
Yukon	395	56	13	14,2	3,3
Territoires du Nord-Ouest	899	437	158	48,6	17,6
Certaines régions métropolitaines de recensement					
Montréal	42 174	12 311	3 296	29,2	7,8
Toronto	40 020	5 473	1 752	13,7	4,4
Vancouver	29 565	6 224	2 133	21,1	7,2

(1) Ne s'applique pas au Nunavut.

Dans l'ensemble du pays, les logements ne faisant pas partie du parc de logements représentent 23,3 % de tous les logements classés comme étant inoccupés lors du recensement. Le problème est légèrement plus marqué dans les régions rurales (25,6 %) que dans les régions urbaines (20,5 %). À l'échelle des provinces et des territoires, le taux de logements qui ne faisaient pas partie du parc de logements ayant été classés comme logements inoccupés variait de 9,5 % à Terre-Neuve-et-Labrador à 48,6 % dans les Territoires du Nord-Ouest. L'Île-du-Prince-Édouard affichait le taux particulièrement élevé de 41,6 %.

Le tableau 7.3.1E présente l'erreur de classement des logements au recensement de 2001 qui est attribuable au fait de classer comme étant inoccupés des logements qui n'auraient pas dû être dénombrés au recensement. Cette erreur est estimée à 1,68 % de tous les logements privés dénombrés au recensement de 2001. Au niveau des provinces et des territoires, cette erreur varie entre 0,41 % au Yukon et 3,81 % à l'Île-du-Prince-Édouard. Dans le cas des trois plus grandes régions métropolitaines de recensement, l'erreur varie entre 0,33 % à Toronto et 0,84 % à Montréal.

Tableau 7.3.1E Erreur de classification des logements au recensement de 2001, Canada, provinces, territoires et certaines régions métropolitaines de recensement, Étude sur la classification des logements de 2001

	Total des logements	Logements hors du parc de logements			
		Nombre estimé	Erreur-type	Taux (%)	Erreur-type (%)
Canada	12 548 588	210 644	15 264	1,68	0,12
Terre-Neuve-et-Labrador	227 570	3 424	776	1,50	0,34
Île-du-Prince Édouard	55 992	2 132	436	3,81	0,78
Nouvelle-Écosse	403 819	5 748	926	1,42	0,23
Nouveau-Brunswick	313 609	5 406	2 030	1,72	0,65
Québec	3 230 196	73 410	9 422	2,27	0,29
Ontario	4 556 240	52 008	8 298	1,14	0,18
Manitoba	477 085	8 466	2 328	1,77	0,49
Saskatchewan	431 628	13 837	2 584	3,21	0,60
Alberta	1 171 841	23 611	3 734	2,01	0,32
Colombie-Britannique	1 643 969	22 108	6 598	1,34	0,40
Yukon	13 793	56	13	0,41	0,09
Territoires du Nord-Ouest	14 669	437	158	2,98	1,08
Certaines régions métropolitaines de recensement					
Montréal	1 473 491	12 311	3 296	0,84	0,22
Toronto	1 671 087	5 473	1 752	0,33	0,10
Vancouver	786 277	6 224	2 133	0,79	0,27

7.3.2 Logements inscrits comme étant non répondants au recensement de 2001

Le tableau 7.3.2 fait état du nombre de logements non répondants au recensement de 2001 et indique le nombre de personnes ajoutées pour ces logements dans le cadre de l'Étude sur la classification des logements. Pour comparer ces résultats aux résultats qui auraient été obtenus selon l'ancienne méthode d'imputation appliquée aux logements non répondants, voir Dick (2002a et 2002b).

Tableau 7.3.2 Nombre de personnes dans les logements classés comme étant non répondants au recensement de 2001, Canada, provinces et territoires, Étude sur la classification des logements de 2001

	N ^{bre} de logements non répondants au recensement	Logements occupés		Personnes	
		Total	Erreur-type	Total	Erreur-type
Canada	179 788	143 681	2 352	317 587	14 841
Terre-Neuve-et-Labrador	1 431	1 185	67	2 268	179
Île-du-Prince-Édouard	508	392	83	978	206
Nouvelle-Écosse	5 063	3 980	332	8 042	511
Nouveau-Brunswick	3 303	2 676	60	5 501	66
Québec	65 787	50 834	1 473	100 741	7 789
Ontario	59 784	48 396	1 686	124 825	12 282
Manitoba	3 798	3 254	142	6 602	161
Saskatchewan	3 246	2 313	144	5 142	172
Alberta	14 197	11 834	370	26 982	423
Colombie-Britannique	22 472	18 697	470	36 269	2 862
Yukon	118	90	10	177	12
Territoires du Nord-Ouest	81	30	21	60	29

8. Contre-vérification des dossiers

8.1 Introduction

Après chaque recensement depuis 1966, la Contre-vérification des dossiers (CVD) a permis d'évaluer le sous-dénombrement brut au recensement, c'est-à-dire d'estimer le nombre de personnes et de ménages omis lors du recensement. Les résultats de la CVD sont combinés à ceux des autres études sur la couverture pour calculer le sous-dénombrement net.

Le sous-dénombrement de la population et des ménages est généralement considéré comme l'une des plus importantes sources d'erreurs touchant les données du recensement. Il introduit un biais négatif dans la mesure où les chiffres publiés par le recensement sous-estiment les vrais totaux de la population et des ménages. Le surdénombrement par contre, introduit un biais positif car les données du recensement surestiment alors les vrais totaux de la population et des ménages. Ces deux sources d'erreur peuvent également fausser les répartitions des caractéristiques de la population et des ménages estimées à partir des données du recensement si les personnes surdénombrées et omises ne possèdent pas les mêmes caractéristiques que les personnes dénombrées.

Les principaux objectifs de la CVD de 2001 étaient :

- a) d'étudier les effets du sous-dénombrement de la population lors du recensement de 2001 et de procéder à l'estimation du sous-dénombrement dans les provinces et territoires ainsi que pour certains sous-groupes importants de la population;
- b) d'étudier les caractéristiques des personnes omises en vue de connaître les causes possibles des erreurs;
- c) d'obtenir une indication du niveau de surdénombrement des personnes vivant dans les logements privés ou collectifs le jour du recensement de 2001.

L'établissement d'estimations du sous-dénombrement et du surdénombrement des familles de recensement, des familles économiques et des ménages ne faisait pas partie du champ initial des études sur la couverture du recensement de 2001. C'est pourquoi on ne dispose d'aucune donnée sur l'erreur de couverture pour ces populations. Au recensement de 1996, le sous-dénombrement des ménages était estimé à 2,49 % pour l'ensemble du pays et le taux de surdénombrement à 0,30 %, pour un taux de sous-dénombrement net de 2,19 %. Pour connaître les statistiques détaillées les plus récentes sur l'erreur de couverture des ménages, le lecteur peut consulter le tableau 10.3 du rapport technique du recensement de 1996 intitulé *Couverture* (n° 92-370-XPB au catalogue).

8.2 Méthodologie – Préparation de l'échantillon

La population cible, qui comprenait toutes les personnes qui auraient dû être dénombrées au recensement de 2001, a été formée à partir de six sources ou bases de sondage. Les cinq premières bases ont été utilisées pour l'estimation du sous-dénombrement et du surdénombrement dans les dix provinces tandis que les estimations pour les trois territoires ont été calculées grâce aux échantillons de la dernière base seulement. En tout, on a choisi 60 653 personnes réparties parmi les bases. Le tableau 8.2 présente la répartition de l'échantillon selon les bases de sondage.

Tableau 8.2 Bases de sondage, Contre-vérification des dossiers de 2001

Base de sondage	Définition	Taille de l'échantillon (personnes)
1. Recensement de 1996	Toutes les personnes dénombrées lors du recensement de 1996.	44 930
2. Omis	Toutes les personnes de l'échantillon de la CVD de 1996 qui ont été classées comme étant omises. Ces personnes conservent les poids qu'elles avaient lors de la Contre-vérification des dossiers de 1996.	2 223
3. Naissances	Tous les enfants nés entre le 14 mai 1996 et le 14 mai 2001.	2 805
4. Immigrants	Tous les immigrants reçus arrivés au Canada entre le 14 mai 1996 et le 14 mai 2001.	2 630
5. Résidents non permanents	Tous les titulaires de permis de travail ou de séjour pour étudiants, de permis ministériel (y compris les prolongations) ou les revendicateurs du statut de réfugié, qui étaient au Canada le 15 mai 2001.	1 060
6. Fichiers d'assurance-santé	Toutes les personnes inscrites dans les fichiers d'assurance-santé du Yukon, des Territoires du Nord-Ouest et du Nunavut, qui étaient au Canada le 15 mai 2001.	7 005
Total		60 653

L'échantillonnage a été effectué indépendamment à l'intérieur de chaque base. Le plan d'échantillonnage variait d'une base à l'autre selon la nature de la liste utilisée. Les taux d'échantillonnage n'étaient pas uniformes à l'intérieur des bases. On a eu recours à des taux d'échantillonnage plus élevés pour les sous-groupes où l'on prévoyait un sous-dénombrement ou un surdénombrement important ou un taux de dépistage plus faible, ceci dans le but de rendre le plan d'échantillonnage plus efficace.

Dans la base du recensement de 1996, le plan d'échantillonnage est à un degré avec stratification démographique et une répartition optimale basée sur les taux historiques de sous-dénombrement et de surdénombrement (ce qui est nouveau pour 2001), le taux de dépistage historique et la taille de la strate. La population était stratifiée par province de résidence, sexe, âge et état matrimonial. Les personnes dénombrées sur des réserves et établissements indiens et dans des logements collectifs au recensement de 1996 formaient des strates séparées. Un échantillon a alors été sélectionné dans chacune des strates ainsi formées afin d'observer le plus de cas de personnes non dénombrées possible. Les fractions de sondage n'étaient pas les mêmes dans chaque strate. Par exemple, lors de la sélection, tout comme en 1996, on a attribué une probabilité d'être choisi plus grande aux hommes célibataires âgés de 20 à 24 ans en 2001, car on avait observé lors des contre-vérifications des dossiers (CVD) précédentes que le sous-dénombrement et le surdénombrement sont toujours plus importants dans cette strate.

La base des personnes omises est une base de sondage conceptuelle puisqu'il n'existe pas de liste de toutes les personnes non dénombrées au recensement de 1996. L'échantillon de cette base est constitué de tous les cas classés comme des « non dénombrés » lors de la CVD de 1996. L'échantillon n'est pas stratifié comme tel quoiqu'il y ait une stratification implicite car les cas non dénombrés en 1996 proviennent de bases et de strates différentes dans la CVD de 1996.

Pour la base des naissances, on a obtenu une copie des enregistrements de naissances de la statistique de l'état civil, pour la période intercensitaire. La base a ensuite été stratifiée par province de résidence de la mère et année de naissance de l'enfant et un échantillonnage à un degré a été utilisé.

On a construit la base des immigrants grâce aux dossiers de l'immigration de Citoyenneté et Immigration Canada. Cette base a été stratifiée par province. Pour les trois provinces où se dirigent la majorité des immigrants (le Québec, l'Ontario et la Colombie-Britannique), on a ensuite créé des strates selon l'année d'arrivée au Canada, c'est-à-dire une strate pour les immigrants arrivés entre 1996 et 1999, une deuxième pour ceux arrivés en 2000 et finalement, une dernière pour ceux arrivés en 2001. Pour chacune de ces deux bases, on a ensuite utilisé l'échantillonnage à un degré.

La base des résidents non permanents a été créée à l'aide des dossiers de Citoyenneté et Immigration Canada. Cette base a été stratifiée par province. Pour le Québec, l'Ontario et la Colombie-Britannique, on a ensuite créé des strates renfermant d'une part les titulaires d'un permis de séjour pour étudiants ou d'un permis ministériel, d'autre part les titulaires d'un permis de travail, et en dernier lieu, les revendicateurs du statut de réfugié. Dans les autres provinces, où ces populations sont plus petites, il n'était pas possible de stratifier les bases de cette façon. Des échantillons à un degré ont été tirés dans les strates ainsi formées.

Enfin, les bases de sondage des trois territoires ont été formées à partir des fichiers d'assurance-santé respectifs. Elles ont été stratifiées par âge, sexe et région (urbaine ou rurale pour le Yukon et les Territoires du Nord-Ouest, et Iqaluit ou non Iqaluit pour le Nunavut). Des échantillons à un degré ont été tirés dans chaque strate.

Un problème qui se pose avec l'utilisation des bases de sondage multiples est la possibilité qu'une personne soit incluse sur plus d'une base. Par exemple, une personne qui est sur la base des immigrants pouvait être au Canada en mai 1996 avec un permis de travail et aurait été dénombrable par le recensement de 1996. Elle serait donc à la fois sur la base des immigrants, et sur celle du recensement si elle était dénombrée, ou sur celle des personnes omises si elle était non dénombrée. Il est donc important d'identifier tous les cas de chevauchement de base, sinon les estimations produites risquent d'être trop élevées puisque les personnes se trouvent à être comptées deux fois. Dans la mesure du possible, on identifie ces cas lors de la constitution des bases de sondage, mais cela peut également se faire plus tard, à partir de l'information fournie par les répondants.

Un autre problème est qu'aucune des cinq premières bases du sondage (voir le tableau 8.2) ne couvre les personnes émigrées ou hors pays au recensement de 1996 qui sont revenues durant la période intercensitaire (« Canadiens de retour au pays »). On estime à 192 310 le nombre de ces personnes dont on n'a pas estimé le nombre de non dénombrées. Les estimations de l'erreur de couverture ne couvrent pas cette population. Cependant, en 2001, contrairement à 1996, la CVD a aussi considéré comme « Canadiens de retour » les personnes qui sont revenues dans une province après avoir habité dans un territoire (17 150) ou une réserve indienne ou un établissement indien **partiellement dénombrés** en 1996 (15 110) et dénombrés en 2001.

8.3 Méthodologie – Collecte des données

8.3.1 Cadre de la collecte des données

Le personnel du bureau central (BC) à Ottawa a travaillé en étroite collaboration avec le personnel de cinq bureaux régionaux (BR) de Statistique Canada à la collecte des données pour l'étape de l'Enquête de la Contre-vérification des dossiers (CVD) de 2001. Ces bureaux étaient situés à Halifax, Sherbrooke, Toronto, Winnipeg et Vancouver. Pour attribuer les cas échantillonnés aux BR, le BC estimait au mieux l'endroit où la personne choisie (PC) résidait pendant la période de collecte. Une fois qu'un cas était attribué à un BR, il ne pouvait pas être transmis à un autre BR, même s'il était établi que la PC avait déménagé à l'extérieur de la région du BR. Des statistiques sur la charge de travail des BR sont présentées au tableau 8.3.1.

La collecte a été principalement menée selon la méthode d'interview téléphonique assistée par ordinateur (ITAO). L'application utilisée a été élaborée conformément à nombre des normes établies pour toutes les enquêtes ITAO menées à Statistique Canada et comportait divers modules reliés entre eux. Les BR y avaient accès à l'aide de l'interface générale, laquelle permettait aux responsables des BR d'affecter les intervieweurs à des types de cas précis et de produire souvent des rapports sur l'état d'avancement.

Les répondants qui ne voulaient pas prendre part à une interview téléphonique recevaient une version papier du questionnaire de l'Enquête de la CVD de 2001. Les PC devaient remplir et renvoyer ce questionnaire, dont les données étaient ensuite entrées sur clavier dans le système ITAO par un intervieweur.

Le questionnaire de l'Enquête de la CVD servait à obtenir les adresses où la PC pouvait avoir résidé le jour du recensement et au cours du mois de mai 2001. Le nom et les données démographiques de tous les membres du ménage de la PC étaient recueillis. Les PC qui n'habitaient pas à leur domicile habituel en mai 2001 à cause de leurs études ou de leur travail devaient répondre à des questions additionnelles sur leurs adresses. Un ensemble de questions se rapportant à l'" exode des cerveaux " étaient posées à une petite partie de l'échantillon de l'Enquête de la CVD. Les PC qui étaient décédées au moment de l'Enquête de la CVD étaient incluses dans le champ de l'étude, et l'interview était menée par personne interposée. On faisait également appel à des répondants substitués lorsque la PC était âgée de moins de 15 ans ou qu'elle était absente ou difficile à joindre pour toute la période d'enquête. La durée moyenne d'une interview ITAO était de 12 minutes.

Tableau 8.3.1 Charge de travail des bureaux régionaux pour la collecte des données, Enquête de la Contre-vérification des dossiers de 2001

Bureau régional	Taille de l'échantillon sur le terrain	Heures d'interview	Taux d'achèvement (%)	Régions incluses
Halifax	12 294	18 113	95,0	Terre-Neuve-et-Labrador, Île-du-Prince-Édouard, Nouvelle-Écosse, Nouveau-Brunswick
Sherbrooke	9 921	16 134	83,2	Québec, Nunavut
Toronto	11 367	15 448	90,1	Ontario
Winnipeg	15 272	16 693	91,5	Manitoba, Saskatchewan, Alberta, Territoires du Nord-Ouest
Vancouver	9 988	12 017	87,3	Colombie-Britannique, Yukon
Bureau central (non acheminé à un bureau régional)	1 811	Diverses régions à travers le Canada
Total	60 653	78 405	89,8	Canada

... n'ayant pas lieu de figurer

8.3.2 Opérations de collecte

Les données d'entrée pour l'interview téléphonique assistée par ordinateur étaient chargées en plusieurs vagues, en fonction du moment où la préparation des lots d'échantillons était terminée. Le tableau 8.3.2A présente la répartition dans le temps des 58 842 cas envoyés sur le terrain. Les bureaux régionaux (BR) commençaient généralement les interviews dès qu'ils recevaient les nouveaux cas; ils ont tenté de mener

une interview dans tous les cas, à l'exception de 353 cas inclus dans l'échantillon du Nunavut. Étant donné que l'information attendue du Nunavut pour établir la base de sondage a été reçue plus tard que prévu, ces 353 cas ont été inclus dans l'Enquête de suivi de la Contre-vérification des dossiers, qui a été menée plus tard à l'aide d'un questionnaire papier (voir la section 8.3.5).

Avant que les intervieweurs ne commencent le dépistage et les interviews, les personnes choisies (PC) résidant dans les territoires ont reçu une lettre les invitant à téléphoner à un BR pour une interview.

Les 58 842 cas envoyés sur le terrain représentaient 97,1 % du nombre de cas inclus dans le Fichier de contrôle de l'échantillon. La majorité des 1 811 cas n'ayant pas été envoyés sur le terrain étaient des personnes choisies à partir de la base du recensement de 1996 qu'on croyait décédées (N = 1 068). En outre, environ 700 cas de PC dans l'échantillon des naissances de 2000 et 2001 n'étaient pas prêts à temps pour la collecte sur le terrain et ont été traités par le personnel du bureau central (BC). Enfin, 38 cas inclus dans la base des immigrants ont été jugés « hors du champ d'observation » et n'ont pas été envoyés sur le terrain pour cette raison.

Tous les soirs, une fois les interviews terminées, les cinq BR transmettaient les données d'enquête par voie électronique au BC. Au matin, le personnel du BC examinait les rapports de transmission et les données recueillies. Les cas jugés impropres à tout traitement ultérieur étaient réactivés et retournés au BR pour que les données requises soient recueillies.

Afin d'accroître les taux de réponse, les chefs des BR communiquaient avec des ministères et organismes gouvernementaux afin d'obtenir des adresses et numéros de téléphone dans les cas où aucun contact n'avait pu être établi. De plus, le BC fournissait aux BR des pistes d'adresses tirées de divers fichiers administratifs lors du processus de collecte des données. Des efforts soutenus visant à interviewer les personnes ayant d'abord refusé de participer à l'enquête ont également permis d'améliorer les taux de réponse.

Tableau 8.3.2A Début de la collecte sur le terrain de l'Enquête de la Contre vérification des dossiers, selon la date, la base de sondage et le nombre de cas

Période ou date	Base de sondage	Nombre de cas	Part non écoulee en jours de la période consacrée aux interviews
9 octobre 2001	Recensement : tous les cas	43 863	193
	Naissances : 1996 à 1998	1 554	
	Immigrants : 1996 à 2000	2 075	
	Omises : tous les cas	2 190	
28 novembre 2001	Naissances : 1999	554	145
8 au 14 janvier 2002	Cas du Yukon et des Territoires du Nord-Ouest	5 170	115-107
7 février 2002	Immigrants : 2001	215	86
	Résidents non permanents : tous les cas	1 057	
	Immigrants : divers (1996 à 2000)	15	
27 février 2002	Cas destinés aux méthodologistes du bureau central (données ITAO ¹ de piètre qualité)	286	67
11 mars 2002	Nunavut : tous les cas	1 835	51

Période ou date	Base de sondage	Nombre de cas	Part non écoulee en jours de la période consacrée aux interviews
13 mars 2002	Cas destinés aux méthodologistes du bureau central (partie II)	28	49
30 avril au 10 mai 2002	Fin de la collecte des données dans les bureaux régionaux		

¹ Interview téléphonique assistée par ordinateur

Tableau 8.3.2B Nombre d'interviews terminées, selon le mois et l'année, Enquête de la Contre-vérification des dossiers de 2001

Mois–Année	oct–01	nov–01	déc–01	jan–02	fév–02	mar–02	avr–02	mai–02	Total
N ^{bre} d'interviews terminées	14 252	13 140	5 633	6 257	5 323	4 279	3 015	964	52 863

Deux rapports de gestion détaillés portant sur l'état d'avancement de l'enquête étaient préparés par le BC. L'un d'eux présentait des statistiques sur les cas se trouvant dans les BR (cas non ouverts, cas terminés et cas ouverts mais non terminés). Le second rapport renfermait des données sommaires sur le nombre d'interviews terminées devant être traitées par le BC (voir la section 8.3.4). Ce rapport sommaire incluait également des projections quant aux cas que les BR auraient à terminer, qui visaient à les aider à atteindre leurs objectifs de collecte.

Le BC vérifiait si les données recueillies sur le terrain étaient complètes et exactes. Les cas étaient rejetés si les données étaient insuffisantes ou si elles avaient été recueillies par erreur pour une personne autre que la PC. Les cas non rejetés étaient mis en lots, lesquels étaient ensuite acheminés à l'équipe du traitement de la Contre-vérification des dossiers au BC.

8.3.3 Dépistage

Le dépistage était un aspect essentiel de l'Enquête de la Contre-vérification des dossiers de 2001. Il s'agit du travail effectué, principalement par le personnel des bureaux régionaux, pour trouver l'adresse et le numéro de téléphone des personnes choisies (PC). Le tableau 8.3.3 montre que 59,8 % (51,9 % + 7,9 %) de tous les cas inclus dans l'Enquête de la CVD ont nécessité un dépistage. Il s'agit des cas où les renseignements initiaux fournis par le système d'interview téléphonique assistée par ordinateur (ITAO) n'ont pas permis de joindre la PC ni un répondant substitut acceptable. Selon ce tableau, dans 19 247 cas, le numéro de téléphone n'était pas fourni par le système ITAO et il a fallu commencer par un dépistage, alors que dans 39 595 cas (23 643 + 11 311 + 4 641), un numéro de téléphone était fourni par le système ITAO. Pour 11 311 cas, le numéro fourni n'a pas permis d'établir un contact, de sorte qu'il a fallu recourir au dépistage par la suite.

Tableau 8.3.3 Aperçu des statistiques sur le dépistage

	Nombre	Pourcentage du total
Cas envoyés aux bureaux régionaux	58 842	100,0
Cas n'ayant nécessité aucun dépistage	23 643	40,2
Cas ayant nécessité un dépistage	30 558	51,9
Dès le début	19 247	32,7
Ultérieurement	11 311	19,2
Cas nécessitant toujours un dépistage à la fin de l'Enquête	4 641	7,9

Aux fins du dépistage, les intervieweurs avaient recours à divers outils, les répertoires électroniques en ligne étant les plus utilisés. Toutefois, les pistes les plus utiles étaient fournies par l'application ITAO elle-même. En outre, le bureau central communiquait aux BR les numéros de téléphone et les adresses qu'il parvenait à obtenir.

8.3.4 Statistiques sur la collecte

Le tableau 8.3.4 présente les taux d'achèvement obtenus dans le cadre de l'Enquête de la Contre-vérification des dossiers de 2001 pour les provinces et les territoires. Dans le cas du Nunavut, le taux s'établirait à 66,0 %, plutôt qu'à 45,1 %, si l'on incluait les cas traités lors de l'Enquête de suivi de la Contre-vérification des dossiers, ce qui porterait le taux d'achèvement national à 90,5 %.

Tableau 8.3.4 Taille de l'échantillon et statistiques sur l'achèvement, Canada, provinces et territoires, Enquête de la Contre-vérification des dossiers

	Taille de l'échantillon	Nombre de cas envoyés sur le terrain	Nombre de cas terminés sur le terrain	Taux d'achèvement (%)
Canada	60 615¹	58 842	52 863	89,8
Terre-Neuve-et-Labrador	3 574	3 449	3 336	96,7
Île-du-Prince-Édouard	2 809	2 716	2 606	95,9
Nouvelle-Écosse	3 501	3 388	3 186	94,0
Nouveau-Brunswick	3 398	3 267	3 027	92,7
Québec	8 370	8 115	7 470	92,1
Ontario	11 268	10 904	9 795	89,8
Manitoba	3 671	3 525	3 232	91,7
Saskatchewan	3 762	3 622	3 434	94,8
Alberta	5 044	4 885	4 345	88,9
Colombie-Britannique	8 213	7 966	6 965	87,4
Yukon	2 065	2 065	1 780	86,2
Territoires du Nord-Ouest	3 105	3 105	2 859	92,1
Nunavut	1 835	1 835	828	45,1

¹ La taille effective de l'échantillon de l'Enquête de la Contre-vérification des dossiers était de 60 615 répondants, mais 38 cas estimés « hors du champ de l'enquête » ont été inclus dans l'échantillon. La taille réelle de l'échantillon de la Contre-vérification des dossiers était donc de 60 653 répondants.

8.3.5 Enquête de suivi de la Contre-vérification des dossiers

L'Enquête de suivi (ES) de la Contre-vérification des dossiers de 2001 avait pour principal objectif de mener un suivi ou une seconde interview auprès d'un sous-ensemble de personnes choisies (PC) lors de l'Enquête de la Contre-vérification des dossiers (CVD). D'une part, l'ES consistait à obtenir d'autres adresses des personnes choisies qui, à la suite de la CVD, ne s'étaient révélées inscrites sur aucun des questionnaires du recensement de 2001. Il s'agissait des cas « non trouvés ». D'autre part, dans les cas de « décès non confirmé », cette enquête visait à confirmer la date du décès de la PC. C'est enfin dans le cadre de l'ES que des interviews ont été menées pour la première fois auprès d'un sous-ensemble de PC du Nunavut. Le tableau 8.3.5.1 donne un aperçu des principales activités liées à cette enquête.

Tableau 8.3.5.1 Calendrier de l'Enquête de suivi de la Contre-vérification des dossiers de 2001

Période ou date	Activité de collecte de l'Enquête de suivi de la Contre-vérification des dossiers
25 mars au 17 mai 2002	Planification de l'ES ¹ et élaboration des systèmes
10 mai 2002	Début de la collecte des données pour les cas de « décès non confirmé » et les cas « non trouvés » (39 % de l'échantillon total de l'ES ¹)
28 mai au 5 juillet 2002	Début de la collecte des données pour le reste des cas « non trouvés » et des cas du Nunavut
28 mai au 15 août 2002	Saisie des données par le personnel du bureau central
31 juillet 2002	Fin de la collecte des données
13 août 2002	Fichier de données prêt pour le traitement au bureau central
25 sept. à déc. 2002	Compte rendu d'enquête, documentation, analyses, atelier fédéral-provincial-territorial, etc.

¹ Enquête de suivi

L'ES de 2001 était une enquête papier et crayon menée par téléphone par le personnel du Bureau régional de Sherbrooke. Aux fins du dépistage, on s'est servi des pistes d'adresse obtenues grâce à l'appariement automatisé des données d'interview téléphonique assistée par ordinateur et des données d'entrée du Fichier de contrôle de l'échantillon avec la base de données du recensement de 2001 (voir la section 8.4). Afin de résoudre les cas restants, le personnel a consulté les adresses incluses dans le fichier d'impôt de 2000 pour le Nunavut et un fichier de numéros d'identification des personnes choisies, classés par nom de municipalité, pour le Yukon. Chaque semaine, le personnel du Bureau régional de Sherbrooke envoyait les cas terminés, ainsi qu'un rapport sur l'état d'avancement de la collecte des données, au bureau central, qui procédait à la saisie des données des questionnaires remplis et utilisait un programme du Système d'analyse statistique pour vérifier la validité et la cohérence des réponses. À partir de ces questionnaires contrôlés, l'équipe de traitement de la CVD a établi un fichier cumulatif en format texte visant à faciliter la classification des PC. Comme chaque questionnaire rejeté au contrôle était évalué individuellement, la vérification de la validité et de la cohérence des réponses en vue de la préparation du fichier destiné à l'équipe du traitement représentait une tâche qui prenait beaucoup de temps.

De précieuses observations et recommandations pour l'ES de 2006 ont été obtenues à partir de deux sources : le rapport de compte rendu des intervieweurs et la séance de compte rendu des bureaux régionaux. Au total, 3 122 heures, soit environ 416 jours-personnes, ont été consacrées à l'ES de 2001 au Bureau régional de Sherbrooke. Cette enquête a obtenu un taux de réponse de 81,4 % (voir les tableaux 8.3.5.2 et 8.3.5.3). Dans 178 (80,2 %) des 222 cas de « décès non confirmé » envoyés sur le terrain, on a obtenu confirmation que la date du décès de la PC était antérieure au 15 mai 2001. Dans 39,5 % des cas « non trouvés » (1 656 cas sur 4 192), on a obtenu une nouvelle piste d'adresse pour la PC. Bien que le taux de réponse de l'ES de 2001 soit inférieur au taux obtenu en 1996 (87,3 %), il

convient de souligner que cette enquête a atteint son principal objectif, qui était d'obtenir des renseignements additionnels sur l'adresse où habitaient les PC le jour du recensement. Pour obtenir une description plus détaillée de l'ES de 2001, voir Armstrong et coll. (2003).

Tableau 8.3.5.2 Taux d'achèvement, cas envoyés et terminés selon le type de cas et la base de sondage, Enquête de suivi de la Contre-vérification des dossiers de 2001

Cas envoyés et terminés selon le type de cas	Base de sondage						Total
	Recensement	Naissances	Immigrants	Omis	RNP ¹	Territoires	
« Non trouvés »							
N ^{bre} de cas envoyés	2 667	108	230	316	234	637	4 192
N ^{bre} de cas terminés	2 258	91	184	248	139	535	3 455
Taux d'achèvement (%)	84,7	84,3	80,0	78,5	59,4	84,0	82,4
« Décès non confirmé »							
N ^{bre} de cas envoyés	217	0	0	5	0	0	222
N ^{bre} de cas terminés	204	0	0	4	0	0	208
Taux d'achèvement (%)	94,0	80,0	93,7
« Nunavut »							
N ^{bre} de cas envoyés	0	0	0	0	0	555	555
N ^{bre} de cas terminés	0	0	0	0	0	383	383
Taux d'achèvement (%)	69,0	69,0
Total des cas envoyés	2 884	108	230	321	234	1 192	4 969
Total des cas terminés	2 462	91	184	252	139	918	4 046
Taux d'achèvement (%)	85,4	84,3	80,0	78,5	59,4	77,0	81,4

... n'ayant pas lieu de figurer

¹ Résidents non permanents

Tableau 8.3.5.3 Taux d'achèvement, Canada, provinces et territoires, Enquête de suivi de la Contre-vérification des dossiers de 2001

	N ^{bre} de cas envoyés	N ^{bre} de cas terminés	Taux d'achèvement (%)
Canada	4 969	4 046	81,4
Terre-Neuve-et-Labrador	244	212	86,9
Île-du-Prince-Édouard	153	131	85,6
Nouvelle-Écosse	227	191	84,1
Nouveau-Brunswick	203	178	87,7
Québec	551	486	88,2
Ontario	898	725	80,7
Manitoba	225	181	80,4
Saskatchewan	251	210	83,7
Alberta	312	259	83,0
Colombie-Britannique	713	555	77,8
Yukon	178	136	76,4
Territoires du Nord-Ouest	377	326	86,5
Nunavut	637	456	71,6

8.4 Méthodologie – Traitement des adresses et classification des personnes choisies

La phase du traitement des données de la Contre-vérification des dossiers (CVD) a pour but de classer chaque personne choisie (PC), c'est-à-dire de déterminer son statut. On utilise ensuite les résultats de cette classification lors du rajustement pour la non-réponse et de l'estimation (voir la section 8.5). Il s'agit principalement de déterminer si la PC faisait partie de la population cible du recensement et si elle a été dénombrée aux adresses recueillies, et de fournir d'autres informations nécessaires pour effectuer le rajustement pour la non-réponse.

Ainsi, pour chaque PC, on vise à déterminer si elle peut ou non être classée en :

- a) personne « listée »;
- b) personne « mobile »;
- c) personne comprise dans la « population cible du recensement »;
- d) personne « dénombrée ».

À titre de renseignement, une PC est « dénombrée » si elle est inscrite sur un questionnaire du recensement (formule 2 ou 3) et qu'il y a un enregistrement correspondant dans la base de données du recensement. La « population cible du recensement » comprend les groupes de personnes énumérées à la section 4.2 du présent document. Une PC est « mobile » si son adresse le jour du recensement est différente des autres adresses qui lui sont associées et qui ont été obtenues indépendamment de la collecte de la CVD. Arbitrairement, les personnes hors cible font partie des « mobiles ». Lorsqu'une PC peut être classée indépendamment de la collecte de la CVD, elle est dite « listée ». Par ailleurs, la classification des « omis » regroupe les personnes non dénombrées de la population cible.

On constate que les PC non mobiles dénombrées sont listées et que les autres PC de la population cible sont non listées.

Les PC pour lesquelles on ne peut déterminer une ou plusieurs des caractéristiques a) à d) sont considérées comme non répondantes. Plus précisément, on dit de la PC qu'elle est :

- non identifiée lorsqu'on n'a pu déterminer si elle était listée ou non;
- non contactée lorsqu'on n'a pu déterminer si elle était mobile ou non;
- non dépistée lorsqu'on n'a pu déterminer si elle était comprise dans la population cible du recensement ou non;
- non classée lorsqu'on n'a pu déterminer si elle était dénombrée ou non.

L'utilité de savoir si une PC est dénombrée va de soi : on veut aussi pouvoir classer les PC selon les caractéristiques a) à c) afin de choisir les répondants les mieux appropriés pour représenter les non-répondants.

Pour la base de sondage du recensement et celle des omis, on constate que :

- toutes les PC non identifiées sont non dépistées;
- toutes les PC non dépistées sont non contactées;
- toutes les PC non contactées sont non classées;
- les PC non classées, au sens restreint (c'est-à-dire non classées, mais contactées), sont comprises dans la population cible, sont mobiles et non listées;
- les PC non contactées, au sens restreint (c'est-à-dire non contactées, mais dépistées), sont comprises dans la population cible et sont non listées;
- les PC non dépistées, au sens restreint (c'est-à-dire non dépistées, mais identifiées), sont non listées.

Pour les autres bases de sondage (excluant celle du recensement et celle des omis), on constate que :

- toutes les PC non identifiées sont non contactées;
- presque toutes les PC non contactées sont non dépistées (exceptionnellement, sept PC non contactées sont dépistées);
- toutes les PC non dépistées sont non classées;
- les PC non classées, au sens restreint (c'est-à-dire non classées, mais dépistées), sont comprises dans la population cible, sont mobiles et non listées;
- les PC non dépistées, au sens restreint (c'est-à-dire non dépistées, mais contactées), sont mobiles et non listées;
- les PC non contactées, au sens restreint (c'est-à-dire non contactées, mais identifiées), sont non listées.

On assigne à chaque unité échantillonnée classée comme étant hors cible la raison de cette classification, par exemple, décès, émigration ou unité déjà représentée par une autre base de sondage.

On détermine aussi l'adresse au jour du recensement de chaque PC classée de la population cible, c'est-à-dire l'adresse où, selon les instructions du recensement, la PC devait être dénombrée. Si une PC de la population cible est dénombrée une seule fois, l'adresse du dénombrement est considérée comme l'adresse du jour du recensement, même si d'autres renseignements soulèvent des doutes quant à la bonne exécution des instructions. Afin de mesurer le surdénombrement, on compte aussi combien de fois les PC classées sont dénombrées et on détermine l'adresse des dénombremens superflus (dénombrements doubles ou triples, et dénombremens de personnes décédées ou émigrées). Pour éviter les doubles comptes, on vérifie aussi si les cas de surdénombrement repérés lors de la CVD sont dans le champ des deux autres enquêtes sur le surdénombrement, à savoir l'Étude par appariement automatisé (ÉAA) et l'Étude sur les logements collectifs (ÉLC). Les cas identifiés comme étant couverts

par l'une ou l'autre de ces deux études sont alors exclus du cadre de la CVD et ne sont pas pris en considération pour la composante des estimations de la CVD portant sur le surdénombrement.

Pour en arriver à ce classement, on consulte les registres de l'état civil (p. ex. dans le cas d'un décès), les renseignements obtenus dans le cadre de l'entrevue de la CVD et de l'enquête de suivi (p. ex. cas se rapportant à l'émigration), les bases de sondage (p. ex. cas de chevauchement de bases), ainsi que d'autres sources d'information telles qu'Internet. Cependant, la majeure partie du travail consiste à effectuer une recherche dans les documents du recensement afin de savoir si la PC est dénombrée à l'une des adresses lui étant associées. Les adresses associées à la PC proviennent de plusieurs sources : de l'adresse de sélection sur la base de sondage, des adresses mises à jour grâce aux dossiers fiscaux, des adresses obtenues lors de l'entrevue de la CVD et de l'enquête de suivi, et des adresses obtenues par appariement avec la base du recensement selon la date de naissance et le sexe de la PC et des membres de son ménage.

Des programmes informatiques mettant à profit des informations comme le code postal de l'adresse à résoudre peuvent parfois à eux seuls permettre de déterminer qu'une adresse est le double d'une autre ou (avec un degré élevé de certitude) que la PC a été dénombrée à cette adresse. En général, le personnel affecté à la recherche ne voit pas les adresses qui sont résolues automatiquement, seulement celles qui ne le sont pas. Pour tenter de résoudre celles-ci, il se sert de formulaires électroniques qui, souvent, contiennent déjà des suggestions de questionnaires du recensement ou de secteurs de dénombrement pouvant correspondre à l'adresse. Ces suggestions proviennent des programmes informatiques mentionnés au début du paragraphe. Parfois le questionnaire du recensement correspondant à l'adresse à résoudre est trouvé parmi ces suggestions, mais sinon, on fait appel à d'autres outils tels que des bottins téléphoniques électroniques. On consigne les résultats des recherches sur les formulaires, qui sont ensuite contrôlés électroniquement afin de minimiser les erreurs.

Un fichier renfermant les résultats des recherches pour chaque adresse est ensuite produit. À partir de ce fichier, on identifie les cas qui exigent une vérification plus approfondie, par exemple les cas de PC trouvée dénombrée à plus d'une adresse, de PC apparemment omise sur un questionnaire, ou de non-réponse. Ce fichier sert également à déterminer quelles unités feront l'objet de l'enquête de suivi (voir la section 8.3.5).

Le tableau 8.4A présente la classification des PC de l'échantillon selon la base de sondage. Des 60 653 PC comprises dans l'échantillon, on constate que 3 164 (2 025 + 1 139) avaient été omises par le recensement de 2001 tandis que 50 521 (46 705 + 3 816) avaient été dénombrées. Afin de ne compter chaque PC qu'une fois, les « non-répondants » comprennent : les PC non identifiées, les PC non dépistées, les PC non contactées et les PC non classifiées, au sens restreint. On remarque que 3 776 PC étaient des non-répondants (329 + 2 899 + 541 + 7), cas qui devaient être compensés à la phase de l'estimation par le rajustement pour la non-réponse (voir la section 8.5).

Tableau 8.4A Classification des personnes choisies selon la base de sondage, Contre-vérification des dossiers de 2001

Classification	Recensement		Naissances		Immigrants		Omis		RNP ¹		Territoires		Total	
	N ^{bre}	%												
Dénombrés, non mobiles (listés)	35 763	79,6	2 535	90,4	1 290	49,0	1 184	53,3	270	25,5	5 663	80,8	46 705	77,0
Dénombrés, mobiles (non listés)	2 970	6,6	29	1,0	390	14,8	209	9,4	137	12,9	81	1,2	3 816	6,3
Omis, non mobiles (non listés)	1 398	3,1	57	2,0	49	1,9	166	7,5	97	9,2	258	3,7	2 025	3,3
Omis, mobiles (non listés)	711	1,6	34	1,2	131	5,0	82	3,7	72	6,8	109	1,6	1 139	1,9
Décédés (mobiles, listés)	1 782	4,0	8	0,3	3	0,1	55	2,5	0	0,0	22	0,3	1 870	3,1
Autres hors cible listés (mobiles)	67	0,1	0	0,0	226	8,6	18	0,8	51	4,8	138	2,0	500	0,8
Émigrés (mobiles, non listés)	271	0,6	6	0,2	21	0,8	78	3,5	51	4,8	2	0,0	429	0,7
Autres hors cible non listés (mobiles)	135	0,3	7	0,2	17	0,6	39	1,8	21	2,0	174	2,5	393	0,6
Non classés	76	0,2	4	0,1	4	0,2	16	0,7	21	2,0	208	3,0	329	0,5
Non dépistés	1 728	3,8	18	0,6	352	13,4	374	16,8	205	19,3	222	3,2	2 899	4,8
Non contactés	22	0,0	107	3,8	147	5,6	2	0,1	135	12,7	128	1,8	541	0,9
Non identifiés	7	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	7	0,0
Total	44 930	100,0	2 805	100,0	2 630	100,0	2 223	100,0	1 060	100,0	7 005	100,0	60 653	100,0

¹ Résidents non permanents

Le tableau 8.4B présente les résultats de la CVD quant au surdénombrement. Le nombre total de cas repérés par la CVD se divise en trois catégories : les cas dans le champ de l'ÉAA, ceux dans le champ de l'ÉLC, et ceux qui ne peuvent être relevés que par la CVD. Ce ne sont que ces derniers cas de surdénombrement qui sont pris en considération dans les estimations globales du surdénombrement. Bien que la CVD permette de mesurer le surdénombrement déjà mesuré par l'ÉAA et par l'ÉLC, les résultats de la CVD sont beaucoup moins précis vu le petit nombre de cas observés.

On peut obtenir des renseignements additionnels sur la classification de la CVD de 2001 en consultant Parenteau (2003).

Tableau 8.4B Nombre de cas de surdénombrement (résultats non pondérés) Canada, provinces et territoires, Contre-vérification des dossiers de 2001

	Dans le champ de l'ÉAA ¹	Dans le champ de l'ÉLC ²	Dans le champ de la CVD ³ seulement	Total
Canada	245	3	221	469
Terre-Neuve-et-Labrador	18	0	10	28
Île-du-Prince-Édouard	8	0	12	20
Nouvelle-Écosse	14	1	15	30
Nouveau-Brunswick	6	1	11	18
Québec	30	0	38	68
Ontario	62	0	42	104
Manitoba	17	1	10	28
Saskatchewan	12	0	17	29
Alberta	16	0	23	39
Colombie-Britannique	37	0	34	71
Yukon	2	0	4	6
Territoires du Nord-Ouest	17	0	5	22
Nunavut	6	0	0	6

¹ Étude par appariement automatisé

² Étude sur les logements collectifs

³ Contre-vérification des dossiers

8.5 Méthodologie – Estimation

Les poids finaux des personnes choisies sont calculés à partir des poids initiaux. Ainsi, le poids initial d'une PC de la base des omis correspond au poids final qui lui avait été attribué lors de la Contre-vérification des dossiers (CVD) précédente (moment où la PC avait été classée comme étant omise). Il convient de noter qu'aux fins du recensement de 2001 les poids finaux de la CVD de 1996 ont été recalculés selon une méthodologie semblable à celle de 2001. Par ailleurs, la base des omis comprend un plus grand nombre de PC que celui identifié en 1996. En effet, selon la nouvelle définition, font partie de la base des « omis » les personnes qui en 1996 avaient été dénombrées sur une formule 4 (Fiche de questionnaire manquant ou incomplet), ou qui avaient été classées comme étant non répondantes mais qui vivaient dans un logement pour lequel une formule 4 avait été établie. Pour les PC sélectionnées dans les autres bases de sondage, les poids initiaux sont généralement égaux à l'inverse des probabilités de sélection dans l'échantillon, sauf pour les PC de la base des résidents non permanents. Pour ces dernières, le poids initial est plus élevé afin de tenir compte des résidents non permanents qui n'étaient pas sur la base de sondage au moment de la sélection de l'échantillon, ce qui était effectivement le cas pour une faible proportion de ces résidents. Les comptes exacts relatifs aux résidents non permanents ont été obtenus plus tard.

Les poids initiaux doivent ensuite être rajustés pour tenir compte de la non-réponse. Le poids des non-répondants est redistribué aux répondants. Dans la mesure du possible, cette redistribution se fait en s'assurant que le poids des non-répondants qui présentent certaines caractéristiques est redistribué parmi des répondants présentant les mêmes caractéristiques. Les caractéristiques prises en compte sont : la strate du plan d'échantillonnage; une indication dans les données fiscales qui permet d'inclure la PC dans la population cible; et, finalement, s'il s'agit d'une PC listée, d'une PC mobile, ou d'une PC faisant partie de la population cible. Effectuée à l'aide du Système généralisé d'estimation (SGE), cette redistribution simule un échantillonnage en plusieurs phases chaque phase correspondant à la

« sélection » d'une catégorie différente de PC : PC identifiée, PC contactée, PC dépistée et PC classifiée. Pour ce qui est de la base de sondage du recensement et celle des omis, l'ensemble des PC contactées constitue un sous-ensemble de l'ensemble des PC dépistées; c'est l'inverse pour les autres bases de sondage. L'ordre utilisé par le SGE varie donc selon la base de sondage. À chaque phase, une caractéristique de plus qu'à la phase précédente est prise en compte. Lorsqu'on ne peut identifier un répondant qui présente les mêmes caractéristiques qu'un non-répondant, on regroupe la strate problématique avec une autre strate jugée semblable.

Traditionnellement, on observe que si aucun rajustement n'est fait, l'estimation du nombre de personnes dénombrées dans les territoires est inférieure au chiffre comparable du recensement, situation qui semble due à une sous-couverture de la base de sondage des territoires. Pour contrebalancer cette sous-couverture, on rajuste les poids des PC sélectionnées dans les territoires par un facteur faisant que l'estimation du nombre de personnes dénombrées égale le chiffre comparable du recensement, et ce, pour chacun des trois territoires.

Une unité qui est répondante pour les fins de la mesure du sous-dénombrement ne l'est pas nécessairement pour les fins de la mesure du surdénombrement. En effet, il se peut qu'une unité soit classée comme étant dénombrée sans qu'une entrevue ait été complétée pour elle dans le cadre de la CVD ou sans que toutes les adresses lui ayant été associées aient été résolues. Les adresses non résolues ou non obtenues pourraient être à l'origine d'un surdénombrement. En conséquence, on applique deux facteurs additionnels de façon à gonfler la variable d'intérêt affichant les surdénombrements. Pour obtenir des renseignements additionnels à ce sujet, on peut consulter Théberge (2002).

Le tableau 8.5A présente les chiffres pondérés de la classification de l'échantillon. Il convient de noter que ces chiffres ne tiennent compte ni des ajustements aux chiffres du recensement (voir le chapitre 7), ni des estimations du surdénombrement, lesquels ont une incidence sur le taux de sous-dénombrement net du recensement.

Tableau 8.5A Classification des personnes choisies (résultats pondérés) selon la base de sondage, Contre-vérification des dossiers de 2001

Classification	Recensement		Naissance		Immigrants		Omisi		RPN ¹		Territoires		Total	
	N ^{bre}	%	N ^{bre}	%	N ^{bre}	%	N ^{bre}	%	N ^{bre}	%	N ^{bre}	%	N ^{bre}	%
Dénombrés, non mobiles (listés)	23 692 419	83.2	1 533 274	90.1	511 639	48.4	633 059	54.4	85 936	27.8	85 832	81.7	26 542 159	80.9
Dénombrés, mobiles (non listés)	1 998 415	7.0	37 434	2.2	303 916	28.8	193 622	16.7	94 168	30.4	4 121	3.9	2 631 676	8.0
Omisi, non mobiles (non listés)	802 590	2.8	40 217	2.4	24 904	2.4	104 258	9.0	35 354	11.4	3 979	3.8	1 011 302	3.1
Omisi, mobiles (non listés)	507 422	1.8	61 73	3.6	98 060	9.3	76 499	6.6	43 987	14.2	4 830	4.6	792 528	2.4
Décédés (mobiles, listés)	1 054 204	3.7	3 912	0.2	1 142	0.1	43 217	3.7	0	0.0	348	0.3	1 102 823	3.4
Autres hors cible listés (mobiles)	36 086	0.1	0	0.0	89 773	8.5	7 185	0.6	14 102	4.6	2 056	2.0	149 202	0.5
Émigrés (mobiles, non listés)	250 194	0.9	13 024	0.8	16 501	1.6	65 233	5.6	23 440	7.6	35	0.0	368 427	1.1
Autres hors cible non listés (mobiles)	123 857	0.4	11 526	0.7	10 488	1.0	39 796	3.4	12 333	4.0	3 885	3.7	201 885	0.6
Total	28 465 187	100.0	1 701 117	100.0	1 056 423	100.0	1 162 869	100.0	309 320	100.0	105 086	100.0	32 800 002	100.0

¹ Résidents non permanents

Les résultats pondérés de la CVD quant au surdénombrement sont présentés dans le tableau 8.5B.

Tableau 8.5B Surdénombrement (résultats pondérés), Canada, provinces et territoires, Contre-vérification des dossiers de 2001

	Dans le champ de l'ÉAA ¹	Dans le champ de l'ÉLC ²	Dans le champ de la CVD ³ seulement	Total de la CVD
Canada	159 502	845	148 813	309 160
Terre-Neuve-et-Labrador	3 844	0	1 615	5 459
Île-du-Prince-Édouard	446	0	805	1 251
Nouvelle-Écosse	4 041	397	4 603	9 041
Nouveau-Brunswick	1 588	58	3 908	5 554
Québec	30 385	0	40 443	70 828
Ontario	76 137	0	49 543	125 680
Manitoba	5 652	390	5 050	11 092
Saskatchewan	3 232	0	6 885	10 117
Alberta	10 503	0	13 416	23 919
Colombie-Britannique	23 303	0	22 385	45 688
Yukon	36	0	79	115
Territoires du Nord-Ouest	248	0	81	329
Nunavut	87	0	0	87

¹ Étude par appariement automatisé

² Étude sur les logements collectifs

³ Contre-vérification des dossiers

Enfin, en ce qui a trait au calcul de la variance des estimations, une innovation importante touchant la CVD de 2001 a été le remplacement de la méthode des groupes aléatoires par des méthodes basées sur le plan d'échantillonnage, grâce au SGE. Selon ces méthodes, le plan d'échantillonnage de la CVD est approximé par un plan stratifié avec probabilités de sélection proportionnelles à la taille. Les tailles retenues sont celles qui reproduisent les poids finaux.

On peut obtenir plus de détails sur les méthodes d'estimation utilisées dans le cadre de la CVD de 2001 en consultant Théberge et Liu (2003).

8.6 Méthodologie – Production des tableaux

En 2001, comme en 1996, on a préparé une série de tableaux qui ont été présentés au Comité fédéral-provincial-territorial de la démographie.

Les estimations de l'erreur de couverture sont présentées dans les quatre séries suivantes :

- Série A : Erreur de couverture de la population selon la province ou le territoire et ventilation démographique
- Série B : Statistiques relatives à la classification des personnes choisies

- Série C : Comparaison entre les estimations du nombre de personnes classées comme étant dénombrées selon la Contre-vérification des dossiers (CVD) et selon le recensement
- Série D : Comparaison entre les estimations du nombre de personnes classées comme étant dénombrées selon la CVD et selon des sources indépendantes (autres que le recensement)

Les statistiques présentées au tableau 5.1 et au chapitre 11 sont tirées de certains tableaux de la série A.

Les estimations de l'erreur de couverture ont été produites à l'aide du Système généralisé d'estimation (SGE) de Statistique Canada et d'une série de programmes écrits en SAS (Système d'analyse statistique). La production d'estimations en parallèle selon ces deux méthodes a permis de garantir la cohérence du calcul. Les estimations des erreurs-types ont été produites au moyen du SGE et du SAS. Des tableaux présentant les estimations de l'erreur de couverture et de l'erreur-type ont été élaborés à l'aide du SAS utilisé conjointement avec Microsoft Excel. Cette méthode de production de tableaux a permis de réduire considérablement le temps nécessaire pour cette tâche par rapport à 1996.

9. Étude par appariement automatisé

9.1 Méthodologie

Le surdénombrement peut survenir de deux façons : la première, et la plus fréquente, est lorsqu'une unité est dénombrée plus d'une fois et la seconde, lorsqu'une unité est dénombrée alors qu'elle ne devrait pas l'être. Au recensement de 2001, le surdénombrement a été mesuré à l'aide de trois études : l'Étude par appariement automatisé (ÉAA), l'Étude sur les logements collectifs (ÉLC) et la Contre-vérification des dossiers (CVD).

L'ÉAA vise à repérer et à estimer le surdénombrement dû à la duplication entre ménages privés différents en cherchant dans la base de données du recensement des paires de ménages situés dans la même région géographique (Atlantique, Québec, Ontario, Ouest canadien comprenant la Colombie-Britannique et les territoires), et comprenant des personnes ayant les mêmes caractéristiques quant au sexe et à la date de naissance complète. Deux ensembles de programmes ont été mis en œuvre. Le premier permettait de repérer des paires de ménages situés dans la même région et comptant au moins deux appariements exacts de personnes (322 298 appariements), alors que le second relevait des paires de ménages situés dans la même circonscription électorale fédérale (CÉF) et comportant un seul appariement exact de personnes (819 481 appariements). On a procédé ensuite à la stratification des paires de ménages repérées. Les questionnaires du recensement d'un échantillon de paires de ménages prélevées dans chacune de ces strates ont été manuellement comparés pour déterminer s'il y avait eu surdénombrement. Plus précisément la liste de personnes inscrites sur le questionnaire d'un des ménages était comparée à la liste des personnes inscrites sur le questionnaire de l'autre ménage. Lorsque les mêmes personnes figuraient sur les deux questionnaires, on considérait qu'il s'agissait d'un cas de surdénombrement.

En 1991, l'opération d'appariement automatisé ne comparait que des ménages résidant dans un même secteur de dénombrement. En 1996, l'ÉAA a été élargie considérablement pour permettre la comparaison des ménages résidant dans une même région géographique. Pour le recensement de 2001, on a introduit une nouvelle opération afin de détecter des paires de ménages résidant dans la région de la capitale nationale et dont un ménage était situé au Québec et l'autre en Ontario. La région de la capitale nationale est une grande zone urbaine qui chevauche deux régions géographiques et se caractérise par une forte mobilité de la population.

9.2 Plan d'échantillonnage et traitement

Les 1 141 779 paires de ménages détectées par les programmes d'appariements constituaient la base de sondage. La méthode de stratification appliquée visait à produire des strates qui seraient à la fois représentatives et de taille raisonnable. Il s'agissait de regrouper des paires de ménages présentant des probabilités similaires de contenir du surdénombrement. La base a été stratifiée en fonction de la province, de la proximité géographique, du nombre d'appariements exacts et de quasi-appariements de personnes entre les ménages et du nombre de personnes dans chaque ménage. Aux fins de l'étude, la proximité géographique se définit comme suit :

- les deux ménages sont situés dans le même secteur de dénombrement (SD);
- les deux ménages sont situés dans la même circonscription électorale fédérale (CÉF), mais dans des SD différents;
- les deux ménages sont situés dans la même province, mais dans des CÉF différentes;
- les deux ménages sont situés dans la même région (Atlantique, Québec, Ontario, Ouest canadien comprenant la Colombie-Britannique et les territoires), mais dans des provinces différentes, ou dans la région de la capitale nationale, l'un étant situé au Québec et l'autre, en Ontario.

D'après l'expérience acquise aux recensements de 1991 et de 1996, on a jugé que les paires dont les ménages présentaient le niveau de similarité le plus élevé contenaient du surdénombrement, sans examen supplémentaire. Un échantillon aléatoire a été ensuite tiré parmi les paires de ménages restantes. On a consulté les copies électroniques des questionnaires des paires de ménages sélectionnées afin de déterminer si les mêmes personnes étaient inscrites sur les deux questionnaires. Sur un échantillon total de 17 275 paires de ménages, on a découvert que 35 224 personnes dans 10 953 paires de ménages étaient inscrites sur deux questionnaires différents. Des poids ont été attribués à ces personnes selon le plan d'échantillonnage. Le tableau 9.2.1 montre le nombre de paires de ménages dans la base de sondage et les estimations finales selon l'ÉAA, alors que le tableau 9.2.2 présente des statistiques relatives à l'échantillon.

Tableau 9.2.1 Nombre de paires de ménages et estimations finales, Canada, provinces et territoires, Étude par appariement automatisé de 2001

	Estimations		
	Nombre de paires de ménages formées	Nombre de paires de ménages contenant du surdénombrement	Nombre de personnes surdénombrées
Canada	1 141 779	62 876	146 412
Terre-Neuve-et-Labrador	9 408	774	1 657
Île-du-Prince-Édouard	1 765	198	439
Nouvelle-Écosse	21 128	1 398	2 875
Nouveau-Brunswick	14 295	1 403	2 608
Québec	303 807	15 829	35 061
Ontario	488 083	21 709	53 378
Manitoba	30 224	1 842	4 128
Saskatchewan	21 684	1 748	3 629
Alberta	100 344	5 868	13 263
Colombie-Britannique	149 486	11 792	28 710
Yukon	402	85	173
Territoires du Nord-Ouest	631	129	325
Nunavut	522	100	165

Tableau 9.2.2 Taille de l'échantillon et résultats du traitement, Canada, provinces et territoires, Étude par appariement automatisé de 2001

	Échantillon		
	Nombre de paires de ménages choisies	Nombre de paires de ménages contenant du surdénombrement	Nombre de personnes surdénombrées
Canada	17 275	10 953	35 224
Terre-Neuve-et-Labrador	423	199	591
Île-du-Prince-Édouard	188	97	268
Nouvelle-Écosse	673	328	954
Nouveau-Brunswick	515	246	670
Québec	3 640	2 355	7 014
Ontario	5 616	3 984	13 159
Manitoba	681	281	885
Saskatchewan	568	234	698
Alberta	1 527	783	2 441
Colombie-Britannique	3 165	2 300	8 116
Yukon	76	41	106
Territoires du Nord-Ouest	118	71	227
Nunavut	85	34	95

10. Étude sur les logements collectifs

10.1 Méthodologie

L'Étude sur les logements collectifs (ÉLC) visait toutes les personnes dénombrées comme résidents habituels (RH) dans un logement collectif non institutionnel et a permis de mesurer le surdénombrement qui se produit lorsqu'un répondant remplit deux questionnaires du recensement à des logements différents. Un logement collectif est un établissement commercial, institutionnel, ou communautaire. Il existe deux genres de logements collectifs : les établissements institutionnels comme les hôpitaux et les prisons, et les logements collectifs non institutionnels comme les hôtels, les YMCA et les résidences scolaires. Contrairement à l'ÉLC de 1996, l'étude de 2001 n'a pas produit d'estimations pour les logements collectifs institutionnels. Ce type de surdénombrement est maintenant couvert par la Contre-vérification des dossiers.

Pendant le dénombrement des logements collectifs le 15 mai 2001, les RH des logements collectifs non institutionnels ont rempli un questionnaire individuel du recensement sur lequel ils devaient inscrire une autre adresse où ils auraient pu être dénombrés. Alors que pour l'ÉLC de 1996 la stratification ne considérait que la province ou le territoire, l'ÉLC de 2001 prenait aussi en compte le type de logement.

Chaque province et territoire s'est vu attribuer un échantillon de taille proportionnelle à la racine carrée de la taille de sa population de RH, avec un minimum de 75 personnes par province ou territoire. Cela assurait une bonne précision dans toutes les provinces et territoires sans qu'il y ait perte majeure de précision au niveau canadien. On a ensuite procédé, à l'intérieur de chaque province et territoire, à la répartition d'un échantillon par type de logement proportionnelle à la taille de sa population de RH, avec un minimum de deux personnes par type de logement. Un échantillon de 4 500 RH a été sélectionné. L'échantillon utilisable était de 4 027 RH car 473 RH n'ont pas été considérés en raison du manque d'information démographique.

10.2 Traitement

Le traitement des données comportait des étapes similaires à celles de la Contre-vérification des dossiers. Pour chaque personne choisie (PC) comprise dans l'échantillon, les questionnaires remplis à l'autre adresse ont été vérifiés afin de déterminer si la PC y avait été dénombrée. Si la PC était inscrite sur un des questionnaires, elle était considérée comme un cas de surdénombrement. Parmi les PC, 347 répondants avaient fourni une autre adresse et 148 PC avaient également été dénombrées dans un logement privé.

Les poids initiaux déterminés de façon inversement proportionnelle à la probabilité de sélection ont été rajustés pour tenir compte des 473 RH qui n'ont pas été considérés en raison du manque d'information démographique. Il s'agissait essentiellement de répartir entre les répondants les poids attribués aux non-répondants.

Le tableau 10.2 montre la répartition de l'échantillon final, le nombre de cas de surdénombrement relevés par l'étude ainsi que le nombre pondéré de personnes surdénombrées.

Tableau 10.2 Taille de l'échantillon, nombre de cas de surdénombrement et nombre pondéré de personnes surdénombrées, Canada, provinces et territoires, Étude sur les logements collectifs de 2001

Logements collectifs non institutionnels			
	Taille de l'échantillon	Cas de surdénombrement	Nombre pondéré de personnes surdénombrées
Canada	4 500	148	2 633
Terre-Neuve-et-Labrador	155	1	31
Île-du-Prince-Édouard	99	1	11
Nouvelle-Écosse	245	5	83
Nouveau-Brunswick	257	37	141
Québec	925	11	575
Ontario	946	27	851
Manitoba	350	5	79
Saskatchewan	205	12	58
Alberta	511	23	405
Colombie-Britannique	549	19	392
Yukon	107	2	7
Territoires du Nord-Ouest	76	1	0
Nunavut	75	4	0

11. Estimations de l'erreur de couverture

11.1 Méthodologie

Soit

$\hat{N}D$ l'estimation du nombre de personnes non dénombrées à leur lieu habituel de résidence selon la Contre-vérification des dossiers (CVD) de 2001, obtenue en additionnant les poids ajustés des personnes comprises dans l'échantillon de la CVD dont le classement final est « non dénombrée au recensement de 2001 »;

R le chiffre officiel du recensement de 2001;

\hat{L} l'estimation du nombre de personnes vivant dans des logements classés comme « logements inoccupés » par le recenseur et l'estimation du nombre de personnes vivant dans des logements classés comme « logements non répondants » par le recenseur, selon l'Étude sur la classification des logements;

$\hat{S}_{ÉAA}$ l'estimation du nombre de personnes surdénombrées selon les résultats de l'Étude par appariement automatisé;

$\hat{S}_{ÉLC}$ l'estimation du nombre de personnes surdénombrées selon les résultats de l'Étude sur les logements collectifs;

\hat{S}_{CVD} l'estimation du nombre de personnes surdénombrées selon les résultats de la CVD, corrigée des effets du surdénombrement relevé par l'Étude par appariement automatisé et par l'Étude sur les logements collectifs;

\hat{S} l'estimation du nombre de personnes dénombrées plus d'une fois ou dénombrées par erreur où

$$\hat{S} = \hat{S}_{ÉAA} + \hat{S}_{ÉLC} + \hat{S}_{CVD}$$

Il est à noter que R inclut \hat{L} c'est-à-dire que, selon le processus décrit au chapitre 7, le nombre estimé de personnes non dénombrées parce qu'elles occupaient des logements classés comme logements inoccupés par le recenseur a été ajouté au chiffre des personnes dénombrées de fait pour produire les chiffres officiels du recensement.

Le taux de sous-dénombrement de la population \hat{T}_O indiquant la proportion de personnes omises au recensement de 2001 par rapport à l'ensemble des personnes qui auraient dû être dénombrées a été calculé comme suit :

$$\hat{T}_O = \frac{\hat{N}D - \hat{L}}{R + (\hat{N}D - \hat{L}) - \hat{S}}$$

Le nombre de personnes qui auraient dû être dénombrées correspond au total de personnes qui ont été dénombrées, $R - \hat{L}$ auquel a été ajouté le nombre de personnes qui auraient dû être dénombrées mais qui ne l'ont pas été, $\hat{N}D$, et duquel a été soustrait le nombre de personnes dénombrées qui n'auraient pas dû l'être, \hat{S} . \hat{L} est soustrait de R parce que le chiffre officiel du recensement inclut le nombre estimé de personnes qui ont été exclues à tort parce qu'elles occupaient des logements classés par erreur comme logements inoccupés. Le taux de sous-dénombrement représente le taux de personnes non incluses dans les chiffres officiels du recensement qui comprend déjà un ajustement pour certaines

catégories de personnes omises. Dans le calcul de \hat{T}_O , on doit soustraire \hat{L} de $\hat{N}D$ parce que ces personnes sont déjà comptées dans R et que $\hat{N}D$ donne une estimation du nombre total de personnes non dénombrées à leur lieu habituel de résidence.

Le taux de surdénombrement de la population \hat{T}_S a été calculé comme suit :

$$\hat{T}_S = \frac{\hat{S}}{R + (\hat{ND} - \hat{L}) - \hat{S}}$$

Le taux de sous-dénombrement net de la population \hat{T}_N a été calculé comme suit :

$$\hat{T}_N = \hat{T}_O - \hat{T}_S = \frac{(\hat{ND} - \hat{L}) - \hat{S}}{R + (\hat{ND} - \hat{L}) - \hat{S}}$$

11.2 Résultats

Les principales estimations des études sur la couverture sont présentés dans les tableaux 11.2.1, 11.2.2 et 11.2.3. Le tableau 11.2.1 contient le nombre de personnes omises, le taux de sous-dénombrement, le nombre de personnes surdénombrées, le taux de surdénombrement, le nombre net de personnes omises et le taux de sous-dénombrement net pour le Canada, les provinces et les territoires. Le tableau 11.2.2 présente les mêmes données selon le sexe et le groupe d'âge. Le tableau 11.2.3 montre le nombre de personnes surdénombrées selon la province ou le territoire de sélection et selon l'étude-source.

11.2.1 Sous-dénombrement brut

Les Territoires du Nord-Ouest affichaient le taux de sous-dénombrement brut de la population le plus élevé (9,10 %) alors qu'à l'échelle des provinces, le taux le plus élevé a été enregistré pour la Colombie-Britannique (5,30 %). Les taux des provinces à l'est de l'Ontario étaient inférieurs au taux national (3,95 %) alors que celui de l'Ontario (4,56 %) le dépassait légèrement. On a observé une variation plus importante du sous-dénombrement selon le sexe et le groupe d'âge. Le sous-dénombrement brut était supérieur chez les hommes (4,90 %) — les femmes ayant affiché un taux de 3,02 % —, et les taux les plus élevés ont été enregistrés chez les jeunes adultes. On a enregistré des taux fort élevés pour le groupe d'âge des 20 à 24 ans (9,85 %), soit 11,68 % pour les hommes et 7,91 % pour les femmes.

11.2.2 Surdénombrement

Le tableau 11.2.3 présente, selon la province ou le territoire, les chiffres de chaque étude sur le surdénombrement qui ont été pris en considération pour estimer le surdénombrement total de la population. L'étude par appariement automatisé (ÉAA) a permis de mesurer le surdénombrement attribuable au fait que des personnes ont été dénombrées dans deux ménages ou plus de la même région (Atlantique, Québec, Ontario, Ouest canadien comprenant la Colombie-Britannique et les territoires) ou dans des ménages formés d'une seule personne à l'intérieur de la même circonscription électorale fédérale. On estime à 146 412 le nombre de personnes tombant dans cette catégorie de surdénombrement. Par ailleurs, l'Étude sur les logements collectifs (ÉCL) visait à évaluer le surdénombrement découlant du fait que des personnes ont été dénombrées à la fois dans un logement collectif et dans un logement privé. Le surdénombrement de ce type a été estimé à 2 633 personnes. Bien que la Contre-vérification des dossiers (CVD) permette de mesurer le surdénombrement provenant de toutes les sources puisqu'on demande aux personnes choisies d'indiquer toutes les autres adresses où elles auraient pu être dénombrées, seul le surdénombrement non mesuré par l'ÉAA ni par l'ÉCL est inclus dans l'estimation du nombre total de personnes surdénombrées. De façon générale, la CVD a servi à mesurer le surdénombrement imputable aux cas suivants :

- personnes dénombrées dans des logements différents;
- personnes dénombrées dans des logements différents mais ayant déclaré des caractéristiques

différentes (sexe et date de naissance);

- personnes dénombrées dans des régions différentes.

On a estimé à 148 811 le nombre de personnes dans cette catégorie de surdénombrement.

Selon l'information du tableau 11.2.1, au niveau national, le coefficient de variation (CV) de l'estimation du surdénombrement total s'élève à 4,9 % (le coefficient de variation est l'erreur-type divisée par l'estimation elle-même exprimé en pourcentage). L'ÉAA est la plus précise des trois études, avec un CV de 2,3 %, tandis que ceux de la CVD et de l'ÉLC s'élèvent respectivement à 9,5 % et 9,4 %. Aux échelons des provinces et des territoires cependant, les CV des estimations de l'ÉAA varient entre 4,7 % et 16,7 %, tandis que ceux des deux autres études sont plus élevés. Globalement, les CV des estimations du surdénombrement total varient de 9,4 % à 21,6 % dans les provinces et territoires.

La Colombie-Britannique affichait le taux de surdénombrement de la population le plus important (1,26 %). Parmi les provinces, les taux de surdénombrement dépassaient le taux national en Saskatchewan (1,06 %), au Québec (1,03 %) et dans les Territoires du Nord-Ouest (1,00 %). Ils étaient inférieurs au taux national au Nunavut (0,59 %), à Terre-Neuve-et-Labrador (0,63 %), au Manitoba (0,80 %) et en Nouvelle-Écosse (0,81 %). Ici encore, on observe des variations plus importantes selon le sexe et les groupes d'âge. Comme c'était le cas pour le sous-dénombrement, les jeunes adultes de 20 à 24 ans affichaient le taux de surdénombrement le plus marqué, ce qui témoigne de leur forte mobilité. Toutefois, contrairement à ce que l'on observait en ce qui touche le sous-dénombrement, le taux de surdénombrement était généralement plus important chez les femmes que chez les hommes. Mis à part le groupe des 20 à 24 ans, le surdénombrement était concentré chez les enfants et les adolescents des deux sexes âgés de 5 à 19 ans. Il s'agissait le plus souvent d'enfants et d'adolescents qui ne vivaient pas avec leurs deux parents ou qui ont été dénombrés plus d'une fois parce que leur famille avait déménagé peu de temps avant ou après le recensement.

11.2.3 Sous-dénombrement net

Les taux de sous-dénombrement net montrent l'incidence nette du sous-dénombrement et du surdénombrement, c'est-à-dire que le nombre de personnes omises est contrebalancé par le nombre de personnes dénombrées par erreur. À l'échelle des provinces, le taux de sous-dénombrement net augmentait d'est en ouest, la Colombie-Britannique ayant affiché le taux le plus élevé (4,04 %). En Ontario, un taux de sous-dénombrement net de 3,68 % dépassait le taux national de 2,99 %, alors qu'au Québec un taux de surdénombrement de 1,03 % contrebalançait un taux de sous-dénombrement de 2,93 %, pour se traduire en un taux de sous-dénombrement net de 1,90 % et inférieur au taux national. Pour ce qui est des territoires, les taux de sous-dénombrement net dépassaient le taux national. Les Territoires du Nord-Ouest affichaient le plus élevé de ceux-ci (8,11 %), suivi du Yukon (4,73 %) et du Nunavut (4,49 %).

Le sous-dénombrement net était fort important chez les jeunes adultes âgés de 20 à 34 ans, et davantage encore chez les hommes de ce groupe d'âge. Alors que pour le groupe plus jeune (20 à 24 ans) un fort taux de surdénombrement a permis de contrebalancer en partie un taux de sous-dénombrement très élevé, pour le groupe d'âge supérieur (25 à 34 ans), le taux de surdénombrement proche du taux national n'a pu compenser que légèrement le taux de sous-dénombrement élevé. Pour ce qui est du sous-dénombrement, le taux de sous-dénombrement net était toujours plus marqué chez les hommes que chez les femmes.

Tableau 11.2.1 Nombre estimés, taux estimés et erreur-type, erreur de couverture de la population au recensement de 2001, Canada, provinces et territoires, Contre-vérification des dossiers de 2001

	Sous-dénombrement de la population				Surdénombrement de la population				Sous-dénombrement net de la population			
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
Canada	1 222 286	42 045	3,95	0,13	297 857	14 525	0,96	0,05	924 429	44 749	2,99	0,14
Terre-Neuve-et-Labrador	12 704	1 689	2,43	0,32	3 303	544	0,63	0,10	9 401	1 782	1,80	0,33
Île-du-Prince-Édouard	2 580	734	1,89	0,53	1 255	245	0,92	0,18	1 325	775	0,97	0,56
Nouvelle-Écosse	32 082	3 955	3,44	0,41	7 561	1 251	0,81	0,14	24 521	4 170	2,63	0,44
Nouveau-Brunswick	26 751	3 229	3,57	0,42	6 656	1 436	0,89	0,19	20 095	3 555	2,68	0,46
Québec	216 312	19 664	2,93	0,26	76 080	7 118	1,03	0,10	140 232	21 033	1,90	0,28
Ontario	540 120	31 601	4,56	0,25	103 771	10 466	0,88	0,09	436 349	33 472	3,68	0,27
Manitoba	40 160	5 115	3,49	0,43	9 257	1 711	0,80	0,15	30 903	5 423	2,69	0,46
Saskatchewan	31 803	3 840	3,18	0,37	10 572	1 954	1,06	0,20	21 231	4 333	2,12	0,42
Alberta	96 941	10 505	3,18	0,33	27 084	4 067	0,89	0,13	69 857	11 308	2,29	0,36
Colombie-Britannique	216 029	14 762	5,30	0,34	51 487	4 834	1,26	0,12	164 542	15 598	4,04	0,37
Yukon	1 683	369	5,59	1,16	260	48	0,86	0,16	1 423	372	4,73	1,18
Territoires du Nord-Ouest	3 700	360	9,10	0,80	405	44	1,00	0,11	3 295	362	8,11	0,82
Nunavut	1 421	410	5,07	1,39	165	27	0,59	0,10	1 256	411	4,49	1,40

Tableau 11.2.2 Nombres estimés, taux estimés et erreur-type, erreur de couverture de la population au recensement de 2001 selon le groupe d'âge et le sexe, Contre-vérification des dossiers de 2001

Groupe d'âge	Sous-dénombrement de la population				Surdénombrement de la population				Sous-dénombrement net de la population			
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
Les deux sexes	1 222 286	42 045	3,95	0,13	297 857	14 525	0,96	0,05	924 429	44 749	2,99	0,14
0 à 4 ans	77 570	13 107	4,42	0,71	16 905	3 165	0,96	0,18	60 665	13 508	3,45	0,74
5 à 14 ans	118 449	15 826	2,90	0,38	61 947	6 050	1,52	0,15	56 502	16 992	1,38	0,41
15 à 19 ans	91 818	11 724	4,36	0,53	38 873	5 371	1,85	0,26	52 945	12 913	2,51	0,60
20 à 24 ans	207 509	14 544	9,85	0,62	56 066	6 558	2,66	0,32	151 443	16 120	7,19	0,71
25 à 34 ans	347 270	16 804	8,07	0,36	39 662	3 866	0,92	0,09	307 608	17 288	7,15	0,37
35 à 44 ans	213 732	18 350	4,04	0,33	25 950	3 227	0,49	0,06	187 782	18 640	3,55	0,34
45 à 54 ans	80 287	13 063	1,79	0,29	17 276	1 776	0,39	0,04	63 011	13 185	1,41	0,29
55 à 64 ans	35 243	10 935	1,22	0,37	10 963	1 540	0,38	0,05	24 280	11 044	0,84	0,38
65 ans et plus	50 408	13 603	1,29	0,34	30 214	8 058	0,77	0,21	20 194	15 833	0,52	0,40
Males	750 245	31 331	4,90	0,19	141 014	8 681	0,92	0,06	609 231	32 634	3,98	0,20
0 à 4 ans	29 978	8 176	3,36	0,89	6 124	640	0,69	0,07	23 854	8 201	2,67	0,89
5 à 14 ans	49 412	10 408	2,38	0,49	33 144	4 367	1,59	0,21	16 268	11 306	0,78	0,54
15 à 19 ans	60 171	9 297	5,49	0,80	15 947	3 319	1,45	0,31	44 224	9 877	4,03	0,86
20 à 24 ans	126 419	11 222	11,68	0,92	26 406	4 735	2,44	0,45	100 013	12 308	9,24	1,03
25 à 34 ans	232 261	13 432	10,67	0,55	22 477	2 905	1,03	0,14	209 784	13 773	9,64	0,57
35 à 44 ans	151 739	14 310	5,71	0,51	12 110	1 512	0,46	0,06	139 629	14 397	5,26	0,51
45 à 54 ans	55 863	9 999	2,50	0,44	7 676	768	0,34	0,03	48 187	10 029	2,16	0,44
55 à 64 ans	19 212	7 819	1,35	0,54	4 759	528	0,33	0,04	14 453	7 837	1,01	0,54
65 ans et plus	25 191	9 095	1,50	0,53	12 371	3 429	0,74	0,21	12 820	9 726	0,77	0,58

Groupe d'âge	Sous-dénombrement de la population				Surdénombrement de la population				Sous-dénombrement net de la population			
	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)	Nombre estimé	Erreur-type	Taux estimé (%)	Erreur-type (%)
Females	472 041	29 016	3,02	0,18	156 842	11 591	1,00	0,08	315 199	31 393	2,02	0,20
0 à 4 ans	47 592	10 421	5,50	1,14	10 781	3 100	1,25	0,36	36 811	10 889	4,26	1,21
5 à 14 ans	69 037	12 046	3,44	0,58	28 803	4 063	1,44	0,21	40 234	12 730	2,01	0,62
15 à 19 ans	31 647	7 188	3,13	0,69	22 925	4 217	2,27	0,43	8 722	8 348	0,86	0,82
20 à 24 ans	81 090	9 346	7,91	0,84	29 660	4 535	2,89	0,46	51 430	10 484	5,02	0,97
25 à 34 ans	115 009	10 240	5,41	0,46	17 186	2 504	0,81	0,12	97 823	10 574	4,60	0,47
35 à 44 ans	61 993	11 573	2,35	0,43	13 84	2 782	0,53	0,11	48 153	11 903	1,83	0,44
45 à 54 ans	24 424	8 422	1,09	0,37	9 600	1 513	0,43	0,07	14 824	8 559	0,66	0,38
55 à 64 ans	16 030	7 649	1,09	0,52	6 204	1 384	0,42	0,09	9 826	7 775	0,67	0,53
65 ans et plus	25 218	10 134	1,13	0,45	17 843	7 269	0,80	0,33	7 375	12 487	0,33	0,56

Tableau 11.2.3 Chiffres tirés de l'Étude par appariement automatisé de 2001, de l'Étude sur les logements collectifs de 2001 et de la Contre-vérification des dossiers de 2001 ayant servi à l'estimation du surdénombrement total de la population, pour le Canada, les provinces et les territoires

Nombre estimé de personnes surdénombrées				
	Étude par appariement automatisé	Étude sur les logements collectifs	Contre-vérification des dossiers (<i>excluant les résultats des autres études</i>)	Total
Canada¹	146 412	2 633	148 811	297 857
Terre-Neuve-et-Labrador	1 657	31	1 615	3 303
Île-du-Prince-Édouard	439	11	805	1 255
Nouvelle-Écosse	2 875	83	4 603	7 561
Nouveau-Brunswick	2 608	141	3 908	6 656
Québec	35 061	575	40 443	76 080
Ontario	53 378	851	49 543	103 771
Manitoba	4 128	79	5 050	9 257
Saskatchewan	3 629	58	6 885	10 572
Alberta	13 263	405	13 416	27 084
Colombie-Britannique	28 710	392	22 385	51 487
Yukon	173	7	79	260
Territoires du Nord-Ouest	325	0	81	405
Nunavut	165	0	0	165

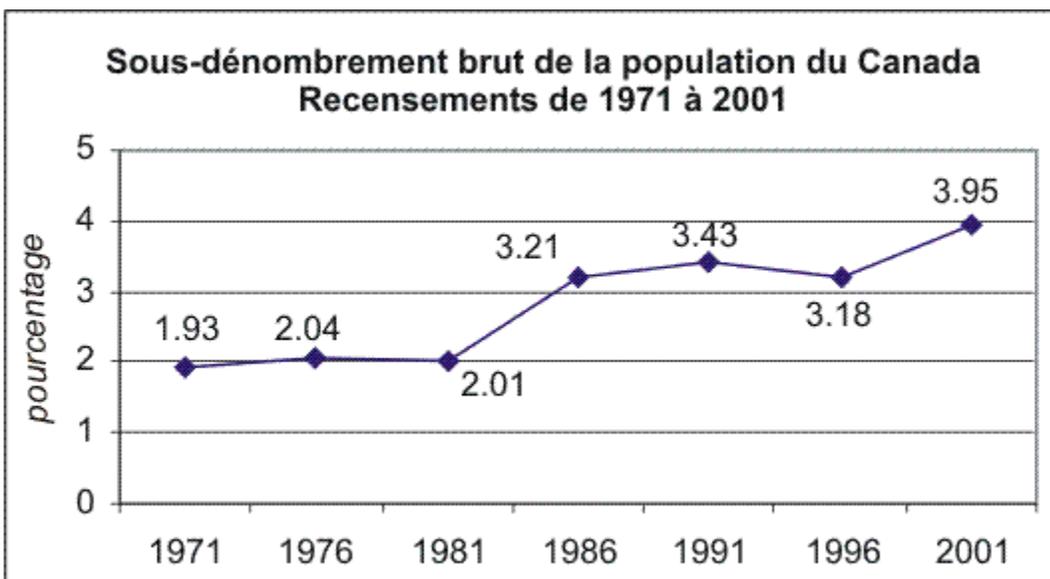
1. L'arrondissement peut être à l'origine de légères différences dans le calcul des totaux pour le Canada.

12. Comparabilité chronologique

12.1 Comparaisons avec les chiffres publiés du recensement de 2001

Dans la présente section, nous établissons des comparaisons entre les taux de sous-dénombrement brut de 1971 à 2001. Les estimations du sous-dénombrement net ne sont disponibles qu'à partir de 1991, puisque c'est lors de ce recensement que l'on a produit pour la première fois une estimation détaillée du surdénombrement, à la suite d'une étude expérimentale réalisée en 1986. Le tableau 12.1A montre les taux de sous-dénombrement brut de la population pour le Canada, les provinces et les territoires de 1971 à 2001, alors que le tableau 12.1B donne les taux selon le groupe d'âge et le sexe.

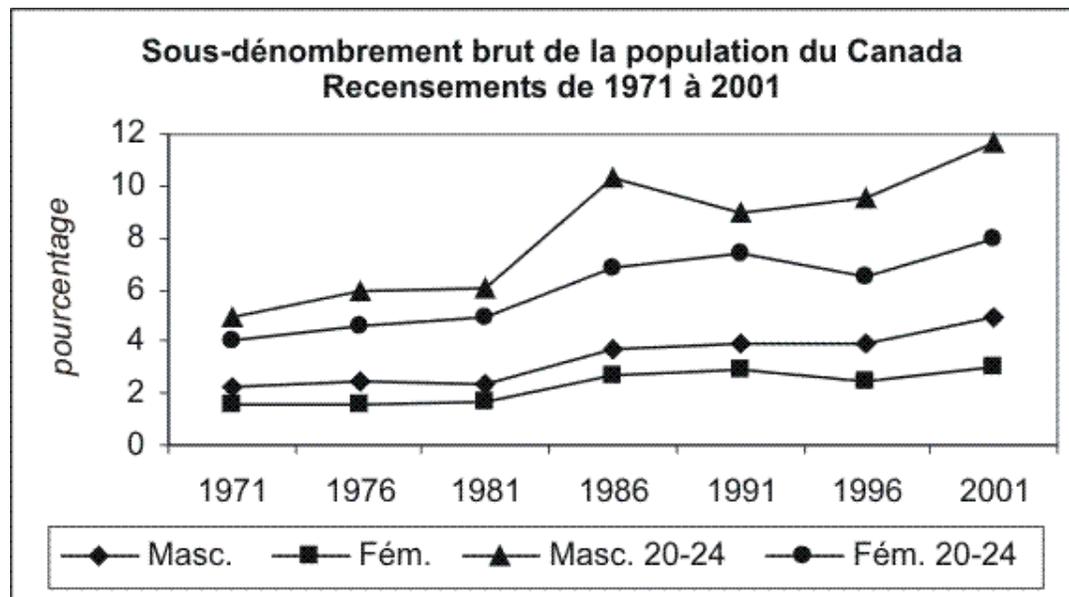
Figure 1. Sous-dénombrement brut de la population du Canada – Recensements de 1971 à 2001



Le taux de sous-dénombrement national brut s'établissait à près de 2 % aux recensements de 1971, 1976 et 1981, puis a atteint 3,21 % en 1986 et 3,95 % en 2001. On pense que les augmentations les plus récentes sont imputables à la fois à l'augmentation du nombre de logements difficiles à dénominer, par exemple les logements rénovés dans les centres-villes, et au changement d'attitude du public à l'égard du gouvernement qui s'est soldé par une baisse du taux de participation au recensement de 2001. À cause de l'augmentation du sous-dénombrement au recensement de 1986, diverses mesures visant à améliorer la couverture ont été prises au recensement de 1991. Ainsi, l'utilisation du Registre des adresses qui constituait une liste distincte des logements devant être dénombrés a permis de maintenir l'erreur de couverture à peu près au même niveau qu'en 1986. Au recensement de 1996, le recours à la méthode de recensement par interview plutôt qu'à la méthode d'autodénombrement, dans certains secteurs de dénombrement des grandes villes, a également contribué à minimiser le sous-dénombrement. De plus, le changement de la date du recensement, qui a eu lieu à la mi-mai plutôt qu'au début juin, a permis de limiter le sous-dénombrement du fait qu'il était plus probable de trouver les gens à la maison et que l'on évitait la période des déménagements. En ce qui touche les études sur la couverture du recensement de 2001, de nouveaux systèmes automatisés ont été utilisés pour recueillir et traiter les données de la Contre-vérification des dossiers (CVD). Même si ces systèmes visaient une plus grande efficacité des opérations, ils ne pouvaient en tant que tels se traduire par des taux accrus d'erreur de couverture. Il semble toutefois qu'ils aient permis une évaluation plus précise de l'erreur de couverture, en raison du nombre plus élevé d'adresses ayant été recueillies et traitées par la CVD de 2001 par rapport aux CVD antérieures.

Les données présentées dans les tableaux 12.1A et 12.1B permettent de relever les nombreuses tendances suivantes :

Figure 2. Sous-dénombrement brut de la population du Canada selon le sexe et une catégorie d'âge – Recensements de 1971 à 2001



a) Parmi les provinces, la Colombie-Britannique affichait le taux de sous-dénombrement brut le plus élevé pour tous les recensements réalisés de 1971 à 2001 sauf celui de 1991 où l'Ontario a enregistré le taux le plus important.

b) Les taux de sous-dénombrement brut des provinces de l'Atlantique et des Prairies sont généralement inférieurs au taux national.

c) On observe deux phénomènes démographiques persistants en ce qui a trait au sous-dénombrement brut. Premièrement, le taux de sous-dénombrement des hommes est supérieur à celui des femmes pour tous les recensements. Deuxièmement, les jeunes adultes âgés de 20 à 24 ans affichent toujours les taux de sous-dénombrement brut les plus élevés.

En raison de légères modifications apportées à la conception des études sur la couverture d'un recensement à l'autre, les taux présentés dans les tableaux 12.1A et 12.1B ne sont pas strictement comparables. Le lecteur doit prendre note des faits suivants :

2001 : ¹

a) La composante institutionnelle de l'Étude sur les logements collectifs a été abandonnée et le surdénombrement de cette population a été estimé par l'Étude de la CVD.

¹ Des estimations du nombre de personnes omises par le recensement de 1996 ont été produites pour refléter autant que possible la méthodologie de la CVD de 2001. Les ajustements apportés aux résultats de 1996 incluent une adaptation aux changements à la stratégie d'ajustement de la non-réponse et à la stratégie de l'estimation (estimer pour toutes les personnes non dénombrées par le recensement après avoir soustrait les imputations des chiffres du recensement). Les chiffres présentés dans cette section ne tiennent pas compte de ces révisions.

b) Comme en 1996, la CVD de 2001 n'a pas estimé le nombre de personnes omises dans les réserves indiennes et établissements indiens partiellement dénombrés. Le nombre de ces réserves et établissements a toutefois diminué en 2001. Par contre, le refus répétitif de certains d'entre eux de participer au recensement rend de plus en plus difficile la tâche d'estimer leur population. La population dans les réserves et établissements partiellement dénombrés a été estimée à environ 34 500 personnes à l'aide d'un modèle statistique (voir la section 14).

c) La CVD a produit des estimations pour toute la population non dénombrée par le recensement, y compris pour les ajustements aux chiffres du recensement avant publication. Ces ajustements, fondés sur les résultats de l'Étude sur la classification des logements, compensaient pour les logements où les répondants n'avaient pu être contactés (formule 4) et pour les logements occupés qui avaient été classés par erreur comme étant inoccupés. Ces ajustements ont été soustraits des chiffres du recensement avant l'estimation du chiffre de la population canadienne et le calcul des taux de sous-dénombrement et de surdénombrement.

1996 :

a) La CVD de 1996 n'a pas estimé le nombre de personnes omises dans les réserves et établissements indiens partiellement dénombrés. En effet, il devient difficile de prélever un échantillon représentatif puisque certaines réserves et certains établissements indiens refusent toujours de participer au recensement. On a estimé, à l'aide d'un modèle statistique, qu'environ 44 000 personnes vivant dans des réserves et établissements indiens partiellement dénombrés ont été omises au recensement de 1996.

b) L'Étude sur les résidents temporaires a été supprimée au recensement de 1996 parce que les données recueillies étaient de qualité douteuse et que la CVD permettait d'obtenir des données de qualité suffisante sur cette forme de sous-dénombrement.

1991 :² Les résidents non permanents ont été inclus dans la population visée pour la première fois en 1991. On a observé chez ce groupe de personnes un taux de sous-dénombrement brut plus élevé que la moyenne. Si ce groupe n'avait pas été inclus dans la population visée au recensement de 1991, on estime que le taux de sous-dénombrement brut pour le Canada aurait été inférieur d'environ 0,3 points au taux enregistré. L'inclusion des résidents non permanents a influé particulièrement sur les taux de sous-dénombrement brut de l'Ontario, de la Colombie-Britannique et du Québec.

² Des estimations révisées du nombre de personnes omises au recensement de 1991 ont été produites par suite de l'analyse des résultats du recensement de 1996. Pour obtenir ces estimations révisées, on a corrigé certains cas qui avaient été classés par erreur comme omis, on a supprimé l'estimation des personnes omises dans les réserves et établissements indiens partiellement dénombrés et on a utilisé des meilleures estimations du nombre de résidents non permanents ainsi qu'une mesure plus adéquate du surdénombrement fondée sur les résultats de 1996. Les chiffres présentés dans cette section ne tiennent pas compte de ces révisions. Voir Tourigny et coll. (1998).

1986 : Les taux présentés ici pour le recensement de 1986 diffèrent des résultats publiés dans le *Guide à l'intention des utilisateurs sur la qualité des données du recensement de 1986 : Couverture*. Les taux figurant dans les tableaux 12.1A et 12.1B comprennent les modifications apportées après la publication du guide de 1986 lorsque les réserves et établissements indiens partiellement dénombrés ont été inclus comme des unités « omises ». Dans la publication initiale de 1986, on les avait considérées comme ayant été « dénombrées » puisque les chiffres du recensement pour les provinces avaient été redressés en fonction d'une estimation des personnes omises dans ces établissements et réserves.

1976, 1971 : Les chiffres des recensements de 1971 et 1976 ne comprennent pas d'estimations des personnes omises dans les logements classés par erreur comme étant inoccupés tirés de la Vérification des logements inoccupés (VLI). Les chiffres des recensements de 1981, 1986, 1991 et 1996 tiennent compte de ces estimations. Si l'on avait inclus les résultats de la VLI de 1976, le taux de sous-dénombrement se serait établi à 1,78 %. Au recensement de 1971, on n'a pas réalisé de VLI.

1971 : Pour 1971, les groupes d'âge au-dessus de 24 ans diffèrent des groupes d'âge montrés pour les autres recensements.

Tableau 12.1A Estimations du sous-dénombrement de la population et de l'erreur-type pour le Canada, les provinces et les territoires, Contre-vérification des dossiers de 1971 à 2001¹

	1971		1976		1981		1986		1991		1996		2001	
	Taux estimé (%)	Erreur- type (%)												
Canada	1,93	0,09	2,04	0,10	2,01	0,09	3,21	0,13	3,43	0,12	3,18	0,09	3,95	0,13
Terre-Neuve-et-Labrador ²	2,25	0,72	1,10	0,39	1,74	0,45	1,92	0,33	2,47	0,30	2,45	0,29	2,43	0,32
Île-du-Prince-Édouard	1,23	1,13	0,38	0,25	1,17	0,54	2,14	0,80	1,67	0,23	1,76	0,28	1,89	0,53
Nouvelle-Écosse	1,33	0,45	0,86	0,34	1,05	0,34	2,15	0,34	2,25	0,36	2,70	0,27	3,44	0,41
Nouveau-Brunswick	1,65	0,56	2,16	-0,37	1,81	0,30	2,71	0,33	3,71	0,42	2,49	0,28	3,57	0,42
Québec	2,10	0,19	2,95	0,25	1,91	0,21	2,91	0,31	3,18	0,20	2,46	0,18	2,93	0,26
Ontario	1,68	0,12	1,52	0,17	1,94	0,14	3,43	0,19	4,23	0,28	3,40	0,18	4,56	0,25
Manitoba	1,13	0,38	1,07	0,33	0,98	0,35	2,94	0,40	2,31	0,36	2,55	0,29	3,49	0,43
Saskatchewan	1,00	0,37	1,33	0,34	0,99	0,37	2,38	0,37	2,15	0,32	3,30	0,32	3,18	0,37
Alberta	2,55	0,44	1,49	0,26	2,54	0,36	3,00	0,32	2,51	0,27	2,99	0,24	3,18	0,33
Colombie-Britannique	2,89	0,39	3,13	0,31	3,16	0,33	4,48	0,36	3,42	0,24	4,58	0,24	5,30	0,34
Yukon	4,12	0,58	3,92	0,51	5,59	1,16
Territoires du Nord-Ouest ³	5,73	0,57	5,68	0,47	9,10	0,80
Nunavut	5,07	1,39

... indisponible

¹ À l'exclusion des réserves et établissements indiens partiellement dénombrés. Les chiffres de 1991, 1996 et 2001 comprennent les résidents non permanents et les territoires. Les chiffres tiennent compte des modifications apportées à la publication initiale de 1986. Les chiffres de 1971 et de 1976 ne comprennent pas d'estimations des personnes omises dans les logements classés erronément comme étant inoccupés.

² Le nom officiel de la province est devenu Terre-Neuve-et-Labrador le 6 décembre 2001.

³ Avant le recensement de 2001, le Nunavut faisait partie des Territoires du Nord-Ouest.

Tableau 12.1B Estimations du sous-dénombrement de la population et de l'erreur-type selon le groupe d'âge et le sexe, Contre-vérification des dossiers de 1971 à 2001¹

Groupe d'âge	1971 ²		1976		1981		1986		1991		1996		2001	
	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)
Les deux sexes, tous les âges	1,93	0,09	2,04	0,10	2,01	0,09	3,21	0,13	3,43	0,12	3,18	0,09	3,95	0,13
0 à 4 ans	1,99	0,27	2,31	0,28	1,21	0,22	2,14	0,49	3,55	0,49	2,89	0,36	4,42	0,71
5 à 14 ans	0,90	0,13	1,20	0,16	1,23	0,21	2,08	0,26	2,49	0,27	1,45	0,14	2,90	0,38
15 à 19 ans	2,60	0,28	1,99	0,38	2,96	0,52	3,58	0,60	3,75	0,42	3,48	0,42	4,36	0,53
20 à 24 ans	4,49	0,28	5,31	0,38	5,51	0,29	8,66	0,46	8,18	0,52	8,00	0,34	9,85	0,62
25 à 34 ans	2,50	0,20	2,85	0,28	2,31	0,28	4,51	0,35	5,65	0,35	5,81	0,29	8,07	0,36
35 à 44 ans	1,54	0,26	2,20	0,26	2,32	0,31	2,84	0,29	2,78	0,24	4,04	0,33
45 à 54 ans	1,40	0,15	1,22	0,33	0,81	0,23	1,58	0,29	1,61	0,27	1,90	0,21	1,79	0,29
55 à 64 ans	1,22	0,18	0,92	0,20	0,91	0,29	2,06	0,31	1,69	0,28	2,23	0,34	1,22	0,37
65 ans et plus	1,20	0,25	0,71	0,30	1,76	0,31	1,51	0,28	1,52	0,26	1,29	0,34
Hommes, tous les âges	2,27	0,12	2,46	0,17	2,37	0,13	3,75	0,16	3,95	0,16	3,89	0,14	4,90	0,19
0 à 4 ans	1,73	0,34	2,53	0,46	1,32	0,33	2,22	0,67	2,79	0,58	2,56	0,47	3,36	0,89
5 à 14 ans	0,93	0,18	1,14	0,21	1,27	0,29	1,98	0,32	2,32	0,34	1,46	0,24	2,38	0,49
15 à 19 ans	2,71	0,39	1,93	0,48	3,12	0,68	4,09	0,74	3,55	0,60	3,68	0,43	5,49	0,80
20 à 24 ans	4,97	0,40	5,99	0,52	6,03	0,48	10,36	0,57	8,98	0,81	9,48	0,50	11,68	0,92
25 à 34 ans	3,38	0,31	3,64	0,46	2,70	0,44	5,43	0,45	7,28	0,56	7,74	0,42	10,67	0,55
35 à 44 ans	2,33	0,48	3,42	0,40	3,29	0,51	3,65	0,41	3,94	0,39	5,71	0,51
45 à 54 ans	1,90	0,24	1,63	0,41	1,21	0,38	1,95	0,52	2,05	0,45	2,12	0,27	2,50	0,44
55 à 64 ans	1,37	0,28	1,28	0,34	0,91	0,40	1,88	0,47	2,04	0,44	2,50	0,54	1,35	0,54
65 ans et plus	1,90	0,44	0,69	0,47	1,57	0,50	1,41	0,50	1,64	0,45	1,50	0,53

Groupe d'âge	1971 ²		1976		1981		1986		1991		1996		2001	
	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)	Taux estimé (%)	Erreur-type ³ (%)
Femmes, tous les âges	1,59	0,11	1,61	0,10	1,65	0,12	2,68	0,17	2,93	0,17	2,49	0,12	3,02	0,18
0 à 4 ans	2,25	0,40	2,07	0,36	1,10	0,33	2,06	0,62	4,35	0,71	3,24	0,55	5,50	1,14
5 à 14 ans	0,87	0,17	1,26	0,27	1,19	0,31	2,20	0,33	2,65	0,39	1,45	0,22	3,44	0,58
15 à 19 ans	2,49	0,38	2,05	0,51	2,80	0,73	3,05	0,76	3,96	0,54	3,28	0,55	3,13	0,69
20 à 24 ans	4,01	0,37	4,62	0,48	4,98	0,43	6,89	0,72	7,36	0,71	6,45	0,48	7,91	0,84
25 à 34 ans	1,58	0,22	2,03	0,38	1,92	0,32	3,59	0,45	3,98	0,37	3,84	0,40	5,41	0,46
35 à 44 ans	0,72	0,24	0,93	0,31	1,33	0,32	2,01	0,35	1,62	0,28	2,35	0,43
45 à 54 ans	0,90	0,17	0,81	0,38	0,41	0,26	1,20	0,35	1,16	0,34	1,68	0,33	1,09	0,37
55 à 64 ans	1,10	0,24	0,58	0,25	0,92	0,34	2,23	0,50	1,35	0,33	1,97	0,40	1,09	0,52
65 ans et plus	0,64	0,38	0,71	0,42	1,89	0,44	1,58	0,36	1,43	0,32	1,13	0,45

... n'ayant pas lieu de figurer

1. Les réserves et établissements indiens partiellement dénombrés sont exclus des chiffres de 1996. Les chiffres de 1991, 1996 et 2001 comprennent les résidents non permanents et les territoires. Les chiffres tiennent compte des modifications apportées à la publication initiale de 1986. Les chiffres de 1971 et de 1976 ne comprennent pas d'estimations des personnes omises dans les logements classés erronément comme étant inoccupés.
2. Les groupes d'âge supérieurs à 24 ans sont : 25 à 39 ans, 40 à 59 ans, 60 ans et plus.

13. Évaluation de la contre-vérification des dossiers

13.1 Introduction

Les résultats de la plus vaste étude sur la couverture, soit la Contre-vérification des dossiers (CVD), peuvent être évalués en comparant les estimations selon la CVD avec des données portant sur la même caractéristique mais tirées d'autres sources, comme la base de données du recensement de 2001. Les comparaisons avec les estimations selon la CVD servent à évaluer ces dernières et à quantifier les différences au chapitre des concepts et des mesures.

Malgré quelques différences conceptuelles entre la CVD et le recensement de 2001, l'estimation du nombre de personnes dénombrées au recensement de 2001 selon la CVD peut être comparée au chiffre de la base de données du recensement de 2001. Afin de rendre ces deux chiffres comparables, certains ajustements ont été apportés aux chiffres du recensement avant d'effectuer la comparaison.

Les estimations des composantes de croissance intercensitaire peuvent être comparées aux estimations de la CVD. En particulier, l'estimation du nombre de personnes qui sont décédées entre le recensement de 1996 et celui de 2001 selon la CVD peut être comparée au chiffre tiré des statistiques de l'état civil.

Les chiffres du solde migratoire interprovincial estimé des données provenant de l'Agence des douanes et du revenu du Canada peuvent être comparés aux estimations selon la CVD. Il est toutefois impossible d'établir des comparaisons rigoureuses pour cette caractéristique, car on ne peut apporter des ajustements suffisants aux différences conceptuelles.

Finalement, la CVD (et le recensement) peut produire des estimations des composantes de l'accroissement de la population, qui peuvent être comparées aux composantes comparables des estimations démographiques tirées d'autres sources. Nous présentons ici les chiffres provinciaux du total des composantes de l'accroissement.

13.2 Comparaisons avec les chiffres publiés du recensement de 2001

Étant donné que le plan d'échantillonnage stratifié à un seul degré de la CVD produit des estimateurs sans biais, les écarts entre les estimations selon la CVD et celles selon le recensement sont attribuables à l'erreur d'échantillonnage entachant les estimations selon la CVD, aux différences conceptuelles entre les deux sources et/ou aux biais affectant les deux sources, qui se traduisent par une sous-estimation ou une surestimation systématique de la caractéristique étudiée.

13.2.1 Personnes dénombrées

Les comparaisons établies pour les provinces et l'ensemble du pays sont présentées au tableau 13.2.1, de même que l'erreur-type de l'estimation selon la CVD et la valeur t servant à tester l'hypothèse selon laquelle il n'y a pas de différences entre l'estimation selon la CVD et le chiffre du recensement comparable. Les ajustements ci-après ont été apportés aux chiffres publiés du recensement, afin de tenir compte des différences conceptuelles entre les deux sources :

- Les ajustements faits à partir de l'Étude sur la classification des logements ont été supprimés, car ils sont inclus dans la base de données du recensement mais ne font pas partie de l'estimation du nombre de personnes dénombrées selon la CVD.
- Le chiffre du surdénombrement du recensement de 2001 est soustrait, car la base de données du recensement contient des personnes surdénombrées alors que l'estimation selon la CVD est établie d'après le nombre de personnes ayant été dénombrées au moins une fois.

- Le nombre de personnes qui résidaient à l'extérieur du Canada cinq ans auparavant (à l'exclusion des immigrants et des résidents non permanents) selon le recensement est également soustrait, car la base de la CVD n'inclut pas ces personnes.
- Enfin, le chiffre du surdénombrement selon le recensement de 1996 est ajouté, puisqu'il est contenu dans l'estimation selon la CVD en raison des poids initiaux associés à la base de sondage du recensement de 1996 qui n'ont pas été corrigés en fonction de ce surdénombrement.

Tableau 13.2.1 Comparaison entre le nombre estimé de personnes dénombrées selon la Contre-vérification des dossiers de 2001 et le chiffre correspondant du recensement de 2001, pour le Canada, les provinces et les territoires

	Contre-vérification des dossiers		Chiffre du recensement publié	Chiffre du recensement comparable ¹	Écart	Valeur <i>t</i> ²
	Nombre estimé	Erreur-type				
Total provincial	29 173 838	57 922	30 007 094	29 152 400	21 438	0,37
Terre-Neuve-et-Labrador	510 247	5 712	512 930	507 577	2 670	0,47
Île-du-Prince-Édouard	129 299	1 500	135 294	132 740	-3 441	-2,29
Nouvelle-Écosse	908 030	10 668	908 007	886 503	21 527	2,02
Nouveau-Brunswick	706 494	6 588	729 498	715 450	-8 956	-1,36
Québec	7 080 809	30 439	7 237 479	7 038 074	42 735	1,40
Ontario	11 042 521	47 884	11 410 046	11 074 981	-32 460	-0,68
Manitoba	1 092 493	10 440	1 119 583	1 102 426	-9 933	-0,95
Saskatchewan	963 583	9 644	978 933	958 663	4 920	0,51
Alberta	2 889 840	21 291	2 974 807	2 883 176	6 664	0,31
Colombie-Britannique	3 760 570	24 448	3 907 738	3 762 859	-2 289	-0,09
Yukon	27 599	0	28 674	27 599	0	...
Northwest Territories	36 487	0	37 360	36 487	0	...
Nunavut	25 866	0	26 745	25 866	0	...

... n'ayant pas lieu de figurer

¹ Le chiffre du recensement comparable est calculé de la façon suivante : Chiffre du recensement comparable = Chiffre du recensement publié – Étude sur la classification des logements – surdénombrement de 2001 – à l'extérieur du Canada en 1996 + surdénombrement de 1996.

² Une valeur *t* supérieure à 1,96 ou inférieure à -1,96 signifie que l'écart est significatif au seuil de 95 %.

Pour l'ensemble du pays, l'estimation des personnes dénombrées au recensement de 2001 selon la CVD est légèrement supérieure (0,07 %) au chiffre comparable du recensement de 2001. Au recensement de 1996 la CVD sous-estimait de 0,08 % le chiffre du recensement et en 1991 elle le sous-estimait de 0,46 % le chiffre du recensement. Pour ce qui est des provinces, l'écart observé pour l'Île-du-Prince-Édouard est statistiquement significatif au seuil de 95 % (-2,29) et celui de la Nouvelle-Écosse est à la limite avec une valeur de 2,02. Les autres écarts ont de valeur-*t* inférieures à 1,5 et on observe six écarts inférieurs à 1,0 (pour Terre-Neuve-et-Labrador, l'Ontario, le Manitoba, la Saskatchewan, l'Alberta et la Colombie-Britannique).

Les écarts importants sont préoccupants, car ils peuvent découler d'un biais dans la classification de la CVD (y compris, par exemple, la province de résidence le jour du recensement). Cependant, d'autres facteurs jouent un rôle important. Outre l'erreur d'échantillonnage, des biais dans les ajustements apportés au chiffre du recensement publié afin d'obtenir un chiffre équivalent du point de vue conceptuel

peuvent expliquer l'écart observé. De plus, le biais de non-réponse de la CVD peut influencer cet écart, car l'ajustement en fonction de la non-réponse est choisi afin d'obtenir le meilleur résultat possible pour estimer les personnes omises et non pour les personnes dénombrées.

Le taux de l'île-du-Prince-Édouard est sensible aux migrations interprovinciales car les personnes sélectionnées à l'extérieur de cette province ont des poids d'échantillonnage plus élevés que les personnes sélectionnées à l'intérieur de la province et la présence ou l'absence de migration interprovinciale dans l'échantillon exerce une influence certaine sur les estimations. En ce qui concerne la Nouvelle-Écosse, on a noté la présence de plusieurs personnes qui ont été sélectionnées dans d'autres provinces et qui ont migré dans cette province.

13.3 Comparaison avec les estimations de la population

13.3.1 Personnes décédées

Le tableau 13.3.1 permet de comparer l'estimation du nombre de personnes décédées pendant la période intercensitaire selon la province de classification de la Contre-vérification des dossiers (CVD) avec les chiffres tirés des statistiques de l'état civil. Les estimations selon la CVD sont plus élevées à l'échelle nationale que les statistiques de l'état civil, l'écart de 16 764 se traduisant par une différence de 1,5 %. Les différences relatives les plus élevées sont observées au Nouveau-Brunswick (3 811 contre 30 388, soit une différence de 12,5 %) et en Saskatchewan (4 624 contre 44 520, ou une différence de 10,4 %). Quant à la valeur *t*, les différences les plus élevées sont observées au Nouveau-Brunswick (1,34) ainsi qu'en Ontario et en Saskatchewan (1,15 respectivement). Toutes les estimations sont donc bien à l'intérieur de l'intervalle de confiance au niveau 95 % et les résultats observés ne soulèvent pas d'inquiétude au sujet de la CVD.

Tableau 13.3.1 Comparaison du nombre de personnes décédées entre le jour du recensement de 1996 et le jour du recensement de 2001, par province de classification, Contre-vérification des dossiers de 2001 et Statistiques de l'état civil

	Contre-vérification des dossiers			Statistique s de l'état civil	Écart	Valeur <i>t</i> ¹
	Taille de l'échantillon	Nombre estimé	Erreur-type			
Canada	1 848	1 102 475	31 195	1 085 711	16 764	0,54
Terre-Neuve-et-Labrador	134	21 794	1 936	21 068	726	0,37
Île-du-Prince-Édouard	109	5 489	530	5 815	-326	-0,61
Nouvelle-Écosse	129	37 534	3 353	39 389	-1 855	-0,55
Nouveau-Brunswick	147	34 199	2 852	30 388	3 811	1,34
Québec	270	253 278	15 825	269 756	-16 478	-1,04
Ontario	367	429 043	22 957	402 707	26 336	1,15
Manitoba	139	48 092	4 255	48 690	-598	-0,14
Saskatchewan	155	49 144	4 032	44 520	4 624	1,15
Alberta	139	85 654	7 643	84 724	930	0,12
Colombie-Britannique	259	138 249	9 048	138 654	-405	-0,04

¹ Une valeur *t* supérieure à 1,96 ou inférieure à -1,96 signifie que l'écart est significatif au seuil de 95 %.

13.3.2 Migration interprovinciale

Le tableau 13.3.2 sert à comparer les estimations du solde migratoire interprovincial intercensitaire selon la Contre-vérification des dossiers (CVD) avec les chiffres correspondants dérivés à partir des fichiers de l'Agence des douanes et du revenu du Canada (ADRC). En général, les données de l'immigration et de l'émigration internes ne sont pas comparables car la CVD ne prend en compte que les mouvements de population survenus entre la date de référence de la base de sondage et le jour du recensement de 2001, alors que les estimations dérivées des données de l'ADRC prennent en compte les mouvements annuels. Ainsi, ne sont présentées ci-dessous que les estimations du solde migratoire.

L'écart observé est significatif dans le cas de l'Île-du-Prince-Édouard (valeur t de -3,72). La CVD estime un solde migratoire négatif alors que les estimations obtenues à partir des données de l'ADRC indiquent un faible solde positif. Pour toutes les autres provinces à l'exception de la Nouvelle-Écosse, les deux séries d'estimations indiquent des soldes migratoires dans la même direction. Pour ce qui est de la Nouvelle-Écosse, l'estimation selon la CVD indique un solde migratoire positif de 5 047 alors que les données administratives de l'ADRC indiquent un solde migratoire négatif de -5 678. La répartition des valeurs t pour cette caractéristique ne porte pas à croire que l'on doive remettre en question les résultats dérivés de la CVD.

Tableau 13.3.2 Comparaison entre le nombre estimé du solde migratoire interprovincial selon la Contre-vérification des dossiers¹ de 2001 et le chiffre correspondant dérivé des données de l'Agence des douanes et du revenu du Canada, selon les provinces

Province	Contre-vérification des dossiers ¹			Chiffre de l'ADRC ²	Écart	Valeur t ³
	Taille de l'échantillon	Nombre estimé	Erreur-type			
Terre-Neuve-et-Labrador	384	-39 705	5 033	-32 502	-7 203	-1,43
Île-du-Prince-Édouard	210	-4 368	1 268	344	-4 712	-3,72
Nouvelle-Écosse	420	5 047	9 255	-5 678	10 725	1,16
Nouveau-Brunswick	280	-17 813	5 050	-8 658	-9 155	-1,81
Québec	195	-50 247	12 956	-69 345	19 098	1,47
Ontario	787	58 954	20 146	65 317	-6 363	-0,32
Manitoba	283	-25 870	7 785	-20 696	-5 174	-0,66
Saskatchewan	374	-32 055	7 977	-24 720	-7 335	-0,92
Alberta	851	140 848	15 819	136 764	4 084	0,26
Colombie-Britannique	618	-35 802	15 339	-33 847	-1 955	-0,13

¹ La Contre-vérification des dossiers exclut les personnes qui habitaient un des trois territoires en 1996, mais demeuraient dans une province en 2001.

² Agence des douanes et du revenu du Canada

³ Une valeur t supérieure à 1,96 ou inférieure à -1,96 signifie que l'écart est significatif au seuil de 95 %.

13.4 Composantes de la croissance démographique

Une comparaison approfondie entre les estimations des composantes de la croissance démographique intercensitaire selon la Contre-vérification des dossiers (CVD) et les estimations de la population établies à partir de sources de données administratives a été effectuée par les membres de la Division de la démographie. Les estimations des composantes de la croissance démographique selon la CVD sont un

sous-produit de la CVD et par conséquent ne sont pas nécessairement très précises. Les estimations de l'accroissement total selon la CVD (et le recensement) de 2001 et selon les sources de données administratives sont présentées au tableau 13.4.

Ces estimations sont formées de nombreuses autres estimations de la croissance démographique (portant sur les naissances, les décès, le nombre d'immigrants, la migration interne, l'émigration, le solde des résidents non permanents et l'accroissement des réserves et établissements indiens non dénombrés), qui sont sujettes à des erreurs plus ou moins grandes en fonction de la source de données. Prenons, à titre d'exemple, l'estimation du solde des résidents non permanents. Il faut également garder à l'esprit que la CVD n'a pas été conçue afin de produire ce genre d'estimations et qu'il s'agit plutôt d'un produit secondaire. Il est donc normal d'observer des différences entre les deux séries d'estimations.

Les estimations de l'une et l'autre sources présentent un écart à l'échelle nationale (1,7 %). On observe qu'elles sont très similaires pour l'Alberta et la Colombie-Britannique. Par ailleurs, les écarts les plus grands sont observés pour le Québec (43 057) et l'Ontario (-41 317) dont les signes s'opposent. Si l'on considère ces écarts en pourcentage de l'estimation selon la CVD (22 % et 5,3 % respectivement), la différence paraît beaucoup plus importante pour le Québec.

Tableau 13.4 Comparaison des estimations d'accroissement de la population selon les chiffres du recensement de 2001 (ajustés avec les résultats de la Contre-vérification des dossiers) et les statistiques de l'état civil, pour l'ensemble des provinces et par province

Province	Contre-vérification des dossiers	Statistiques de l'état civil	Écart
	Nombre estimé	Nombre estimé	
Total provincial	1 428 406	1 453 268	-24 862
Terre-Neuve-et-Labrador	-35 628	-26 882	-8 746
Île-du-Prince-Édouard	-2 275	2 647	-4 922
Nouvelle-Écosse	23 647	11,154	12 493
Nouveau-Brunswick	-11 570	2 025	-13 595
Québec	191 697	148 640	43 057
Ontario	772 183	813 500	-41 317
Ontario	7 709	16 525	-8 816
Saskatchewan	-13 171	-2 139	-11 032
Alberta	287 507	281 558	5 949
Colombie-Britannique	208 308	206 240	2 068

14. Réserves et établissements Indiens ayant opposé un refus

14.1 Introduction

Lors du recensement de 2001, le dénombrement n'a pas été autorisé ou a été interrompu avant d'être mené à terme dans un certain nombre de réserves indiennes et d'établissements indiens. Ces régions (30 au total) sont désignées par le terme « réserves indiennes et établissements indiens partiellement dénombrés ». On ne dispose donc pas de données du recensement pour ces régions, de sorte qu'elles ne figurent dans aucune totalisation du recensement.

Ces régions occasionnent des problèmes uniques en ce qui touche les études sur la couverture et le Programme des estimations de la population. La population cible de la Contre-vérification des dossiers (CVD) est la même que celle du recensement, et est décrite à la section 4.2 du présent document. Cependant, la population qui est effectivement couverte par la CVD n'inclut pas les résidents pour lesquels des données du recensement de 2001 n'ont pas été recueillies. Aux fins de la CVD, ces résidents seraient considérés comme étant hors du champ d'observation. Toutefois, le Programme des estimations de la population exige une estimation de la population des résidents permanents de ces régions. En conséquence, puisque ni le recensement ni la CVD n'est en mesure de produire une estimation de la population vivant dans ces régions, il a fallu adopter une autre approche, qui est décrite dans le présent chapitre.

Le tableau 14.3.1 présente la liste de toutes les réserves indiennes et établissements indiens qui, selon la définition adoptée pour la collecte des données, ont été « partiellement dénombrés » au recensement de 2001. Il indique également si le dénombrement de ces réserves et établissements était complet ou incomplet au recensement de 1996. Il convient de prendre note que 9 de ces 30 réserves et établissements avaient été complètement dénombrés en 1996, alors que les 21 autres avaient été « partiellement dénombrés » ou avaient « opposé un refus ». Les résultats de 2001 sont très différents de ceux obtenus au recensement de 1996, alors que 77 réserves et établissements comptant environ 50 000 personnes avaient été classés comme étant « partiellement dénombrés » (Hamel, 1997).

14.2 Méthodologie

Afin d'estimer la population des réserves et établissements indiens partiellement dénombrés en 2001, on a élaboré un modèle d'estimation en deux étapes consistant d'abord à construire un modèle de régression linéaire simple permettant d'estimer le chiffre du recensement de 2001, puis à corriger ce chiffre estimé en fonction du sous-dénombrement net qui touche tous les chiffres du recensement, afin qu'il soit conforme aux résultats de la Contre-vérification des dossiers (CVD).

Le modèle de régression linéaire a été construit à partir de toutes les réserves et tous les établissements indiens ayant été complètement dénombrés aux recensements de 1996 et de 2001. Le modèle de base suppose une croissance linéaire de 1996 à 2001 pour toutes les provinces. Cependant, dans le cas de la coordonnée à l'origine et des paramètres de régression, des estimations distinctes ont été produites pour chaque province selon ce modèle simple. Le modèle a été évalué en ce qui touche les hypothèses de régression de base que sont l'indépendance des erreurs, l'homogénéité des variances et la normalité des erreurs. Pour de plus amples renseignements, voir Dick et So (2004).

Comme variable d'entrée pour chaque réserve et établissement indiens partiellement dénombrés en 2001, le modèle de régression utilisait soit le chiffre réel du recensement de 1996, soit le chiffre le plus probable calculé à l'aide du modèle élaboré par Hamel (1997). Le résultat obtenu était le chiffre estimé du recensement de 2001.

Pour que ces chiffres du recensement soient cohérents avec toutes les estimations de la population tirées de la CVD, il faut les corriger en fonction du sous-dénombrement net. Le nombre net de personnes omises a été estimé en calculant le taux de sous-dénombrement net pour toutes les réserves et établissements indiens complètement dénombrés dans chaque province. Ce taux provincial était ensuite appliqué au chiffre estimé « du recensement » pour toutes les réserves et établissements indiens partiellement dénombrés dans la province.

On additionnait ensuite le chiffre estimé « du recensement » et le « nombre estimé net de personnes omises » dans chaque réserve ou établissement afin d'établir la population « estimée » de la réserve indienne ou de l'établissement indien partiellement dénombré.

14.3 Résultats en 2001

Les résultats du modèle de régression sont présentés au tableau 14.3.1 dans la colonne « Chiffre estimé du recensement de 2001 ». Les estimations de 2001 fondées sur le modèle de régression sont supérieures d'environ 10 % à celles de 1996. Le nombre de personnes omises dans chaque réserve ou établissement est estimé en appliquant les taux de sous-dénombrement net aux chiffres estimés au recensement. Le chiffre total définitif de 34 541 correspond au chiffre estimé des résidents des réserves et établissements indiens partiellement dénombrés en 2001.

Tableau 14.3.1 Résultats obtenus pour les réserves indiennes et les établissements indiens partiellement dénombrés en 2001¹ selon la province et le nom de la réserve ou de l'établissement

Province / Nom de la réserve ou de l'établissement	Dénombrement complet/incomplet au recensement de 1996	Chiffre réel ou estimé de 1996	Chiffre estimé du recensement de 2001	Taux de sous- dénombrement ² en 2001 (%)	Nombre estimé de personnes omises en 2001 ³	Estimation de la population en 2001
Québec⁴		10 717	11 717	7,4	931	12 648
Akwesasne (Partie)	Incomplet	2 265	2 471		196	2 667
Doncaster 17	Complet	0	4		0	4
Kahnawake 14	Incomplet	7 087	7 757		617	8 374
Kanesatake	Incomplet	1 137	1 235		98	1 333
Lac-Rapide	Complet	228	250		20	270
Ontario⁴		13 176	14 345	10,1	1 615	15 960
Akwesasne (Part) 59	Incomplet	1 229	1 335		150	1 485
Bear Island 1	Complet	153	169		19	188
Chippewas of the Thames First Nation 42	Incomplet	777	839		94	933
Goulais Bay 15A	Incomplet	46	54		6	60
Marten Falls 65	Complet	204	228		26	254
Moose Factory 68	Complet	0	4		0	4
Munsee-Delaware Nation 1	Incomplet	214	235		26	261
Ojibway Nation of Saugeen (Savant Lake)	Incomplet	203	224		25	249
Oneida 41	Incomplet	1 226	1 332		150	1 482

Province / Nom de la réserve ou de l'établissement	Dénombrement complet/incomplet au recensement de 1996	Chiffre réel ou estimé de 1996	Chiffre estimé du recensement de 2001	Taux de sous- dénombrement ² en 2001 (%)	Nombre estimé de personnes omis en 2001 ³	Estimation de la population en 2001
Pikangikum 14	Complet	1 170	1 261		142	1 403
Rankin Location 15D	Incomplet	670	723		81	804
Six Nations (Part) 40 ⁵	Incomplet	632	681		77	758
Six Nations (Part) 40	Incomplet	4 731	5 174		583	5 757
Tyendinaga Mohawk Territory	Incomplet	1 150	1 249		141	1 390
Wahta Mohawk Territory	Incomplet	154	171		19	190
Whitefish Bay 32A	Incomplet	502	538		61	599
Whitesand	Complet	115	128		14	142
Manitoba⁴		85	97	12,1	13	110
Dakota Tipi 1	Incomplet	85	97		13	110
Saskatchewan⁴		451	491	15,6	90	581
Big Head 124	Incomplet	451	491		90	581
Alberta⁴		3 775	4 115	17,3	862	4 997
Ermineskin 138	Incomplet	1 300	1 413		296	1 709
Little Buffalo	Incomplet	221	243		51	294
Saddle Lake 125	Incomplet	2 254	2 459		515	2 974
Colombie-Britannique⁴		225	255	2,9	8	263
Esquimalt	Incomplet	82	93		3	96
Marble Canyon 3	Complet	67	76		2	78
Pavilion 1	Complet	76	86		3	89
Provincial Total		28 429	31 020	10,4	3 521	34 541

¹ Mise en garde : Les estimations de 2001 présentées dans le tableau 14.3.1 sont purement « fondées sur un modèle ». Leur validité dépend grandement de la mesure dans laquelle le modèle a véritablement saisi la situation telle qu'elle se présentait. Il s'agit d'une hypothèse non vérifiée : ces estimations doivent donc être utilisées avec une extrême prudence.

² Le taux de sous-dénombrement net a été estimé à partir des réserves et établissements indiens complètement dénombrés dans chaque province. Par « réserves et établissements indiens », on entendait les types de subdivision de recensement suivants : réserve indienne, *Indian Government District*, établissement indien, terre inuite, village nordique et village Cri.

³ Estimation du nombre de personnes omises = Chiffre estimé du recensement / (1 - Taux estimé de personnes omises).

⁴ L'arrondissement peut être à l'origine de légères différences dans le calcul des taux.

⁵ Il y a deux inscriptions pour *Six Nations (Part 40)*. La première renvoie au code 3528037 de la Classification géographique type, alors que la deuxième renvoie au code 3529020.

Annexe A. sigles et acronymes

Les sigles et acronymes suivants figurent dans le présent rapport.

ADRC	Agence des douanes et du revenu du Canada
ALENA	Accord de libre-échange nord-américain
AR	agglomération de recensement
BC	bureau central (de Statistique Canada)
BR	bureau régional (de Statistique Canada)
CÉF	circonscription électorale fédérale
CR	commissaire au recensement
CTI	Classification type des industries
CV	coefficient de variation
CVD	Contre-vérification des dossiers
DCR	district de commissaire au recensement
ÉAA	Étude par appariement automatisé
ÉCL	Étude sur la classification des logements
ÉLC	Étude sur les logements collectifs
ES	Enquête de suivi (de la Contre-vérification des dossiers)
ITAO	interview téléphonique assistée par ordinateur
NIM	Nearest-neighbour Imputation Method
PC	personne choisie
RH	résident habituel
RMR	région métropolitaine de recensement
RNP	Résidents non permanents
RV	Registre des visites
SAS	Système d'analyse statistique
SCANCIR	Système canadien de contrôle et d'imputation du recensement
SCIAN	Système de classification des industries de l'Amérique du Nord
SD	secteur de dénombrement
SGE	Système généralisé d'estimation
SPIDER	System for Processing Instructions from Directly Entered Requirements
VLI	Vérification des logements inoccupés

Annexe B. Glossaire

Les définitions des termes, variables et concepts du recensement figurent dans le *Dictionnaire du recensement de 2001* (no 92-378-XIF au catalogue). L'utilisateur doit se reporter au Dictionnaire du recensement de 2001 pour obtenir les définitions complètes et des observations additionnelles se rapportant aux concepts, comme des renseignements sur les variables directes et dérivées, ainsi que leur univers respectif. Les termes suivants ne figurent pas dans le *Dictionnaire du recensement de 2001*, mais s'appliquent dans le cas présent.

Surdénombrement au recensement : Il y a surdénombrement lorsqu'une unité ne faisant pas partie d'un univers du recensement est dénombrée lors du recensement. Le surdénombrement peut survenir de deux façons. La première, et la plus fréquente, est lorsqu'une unité faisant partie d'un univers du recensement est dénombrée plus d'une fois. La seconde est lorsqu'une unité ne faisant pas partie de l'univers du recensement (par exemple un résident étranger, une personne fictive ou un logement marginal inoccupé) est dénombrée par erreur.

Sous-dénombrement au recensement : Il y a sous-dénombrement lorsqu'on omet complètement de dénombrer une unité faisant partie d'un univers visé par le recensement.

Sous-dénombrement net : Le sous-dénombrement net correspond à la différence entre le sous-dénombrement total et le surdénombrement total dans un univers du recensement.

Annexe C. Produits et services du Recensement de 2001

Le recensement constitue une source de données fiables pour décrire les caractéristiques des personnes et des logements au Canada. La gamme de produits et services élaborée à partir des données du recensement est conçue de façon que ces données soient utiles, compréhensibles et accessibles à tous les utilisateurs. D'autres sources, comme le *Catalogue du recensement de 2001*, le site Web de Statistique Canada (www.statcan.ca) et, en particulier, le Catalogue en ligne, contiennent de l'information détaillée sur la gamme complète des produits et services du recensement de 2001.

Les produits et services comportent de nouvelles caractéristiques à l'occasion du recensement de 2001 :

1. Support

- On favorise Internet comme moyen pour diffuser les produits de données normalisés et les produits de référence.
- Davantage de données sont offertes gratuitement au grand public à partir d'Internet.

2. Contenu

- On diffuse les tableaux de données du recensement de 2001 selon des **thèmes**, c'est-à-dire des groupements de variables dont les sujets sont apparentés.
- Dans la mesure du possible, on a simplifié la langue et le vocabulaire dans les produits du recensement de 2001 diffusés dans Internet afin de rendre l'information accessible à un plus grand nombre de personnes.
- On a mis en place des moyens qui permettent aux utilisateurs de faire des recherches et de naviguer dans les **produits normalisés du recensement** (y compris les **produits de référence**) diffusés dans Internet.

3. Géographie

- Des unités géographiques, comme les aires de diffusion, les régions urbaines, les localités désignées et les zones d'influence métropolitaine, s'ajoutent à la gamme des produits normalisés. Certaines des nouvelles unités, comme les aires de diffusion, remplacent d'autres unités.

4. Variables

- Au recensement de 2001, on a recueilli de l'information sur de nouveaux sujets : le lieu de naissance des parents, les autres langues parlées à la maison et la langue de travail. Par ailleurs, la question sur la religion, qui est posée lors de chaque recensement décennal, figurait dans le questionnaire de 2001. En outre, on a élargi la définition de la variable portant sur la structure de la famille afin d'inclure les couples de même sexe.

Bibliographie

ARMSTRONG, M., H. FARR et H. RICHARDS. 2003. Enquête de la Contre-vérification des dossiers 2001 : Rapport de la collecte, Ottawa, Statistique Canada. Document de travail non publié.

CLARK, C., M. ARMSTRONG et C. THIBAUT. 2003. Mesures et innovations dans les études de dénombrement du Recensement de 2001, Ottawa, Statistique Canada. Actes de la réunion annuelle de la Société statistique du Canada, Halifax, Nouvelle-Écosse, 8 au 11 juin 2003.

DICK, P. 2002a. The Census of Canada: The Dwelling Classification Study [Le recensement du Canada : L'Étude sur la classification des logements], Ottawa, Statistique Canada. Titre d'un exposé donné au Joint Statistical Meeting of the American Statistical Society, New York, 11 au 15 août 2002.

DICK, P. 2002b. Recensement du Canada : Étude sur la classification des logements, Ottawa, Statistique Canada. Titre d'un exposé donné à la réunion fédérale-provinciale/territoriale sur la démographie, Ottawa, 12 septembre 2002.

DICK, P., et G. SO. 2004. Estimating the Population Count on Incompletely Enumerated Indian Reserves in the 2001 Census [Estimations des chiffres de population dans les réserves indiennes partiellement dénombrées au recensement de 2001], Ottawa, Statistique Canada. Document de travail non publié.

HAMEL, D. 1997. Rapport des méthodes d'estimation de la population et du nombre de logements privés occupés dans les réserves indiennes partiellement dénombrées au recensement de 1996, Ottawa, Statistique Canada. Document de travail non publié.

MOREL, J., et G. KLEIM. 2003a. Recensement de 2001 : Études de couverture, Rapport sur le dénombrement, Évaluation, Ottawa, Statistique Canada. Document de travail non publié.

MOREL, J., et G. KLEIM. 2003b. Recensement de la population de 2001 : Études de couverture, Rapport sur le dénombrement, Revue de la méthodologie, Ottawa, Statistique Canada. Document de travail non publié.

PARENTEAU, M. 2003. Classification de la Contre-vérification des dossiers de 2001, Ottawa, Statistique Canada. Document de travail non publié.

STATISTIQUE CANADA. 2002a. Dictionnaire du recensement de 2001, produit n° 92-378-XIF au catalogue de Statistique Canada, Ottawa, « série des produits de référence du recensement de 2001 ».

STATISTIQUE CANADA. 2002b. Le recensement de 2001 en bref, produit n° 92-379-XIF au catalogue de Statistique Canada, Ottawa, « série des produits de référence du recensement de 2001 ».

STATISTIQUE CANADA. 1999. Couverture, produit n° 92-370-XPB au catalogue de Statistique Canada, Ottawa, « série des rapports techniques du recensement de 1996 ».

THÉBERGE, A. 2002. Facteurs de rajustement pour la surcouverture dans la CVD. Ottawa, Statistique Canada. Document de travail non publié.

THÉBERGE, A., et W.C. LIU. 2003. Le rajustement pour la non-réponse et l'estimation de la variance pour la Contre-vérification des dossiers de 2001, Ottawa, Statistique Canada. Document de travail non publié.

THIBAUT, C. 2003. Recensement de la population de 2001 : Rapport d'évaluation des études de couverture, Ottawa, Statistique Canada. Document de travail non publié.

TOURIGNY, J., M. BUREAU et C. CLARK. 1998. Estimations directes finales de l'erreur de couverture du recensement de 1991, Ottawa, Statistique Canada. Document de travail non publié.