



Guide de l'utilisateur des microdonnées

ENQUÊTE SUR LES COMMUNAUTÉS

2005



Statistique
Canada

Statistics
Canada

Canada

Table des matières

1.0	Introduction	5
2.0	Contexte	7
3.0	Objectifs	9
4.0	Concepts et définitions	11
4.1	Unité d'analyse	11
4.2	Personne la mieux renseignée et conjoint	11
4.3	Variables dérivées relatives à la famille	12
4.4	Variables dérivées relatives au statut socio-économique	14
4.5	Imputation du revenu	14
4.6	Instrument de mesure du développement du jeune enfant	16
5.0	Méthodologie de l'enquête	17
5.1	Population visée	17
5.2	Plan d'échantillonnage	17
5.3	Taille de l'échantillon	18
6.0	Collecte des données	19
6.1	Processus relatif au consentement éclairé	19
6.2	Collecte des données auprès des ménages	19
6.2.1	Composante « Entrée/Sortie »	19
6.2.2	La composante « Enfant »	20
6.2.3	La composante « Adulte » (Parent)	20
6.3	Mesures cognitives recueillies auprès de l'enfant	20
6.4	Durée de l'interview	21
6.5	Formation, supervision et contrôle relativement à l'interview	21
7.0	Traitement des données	23
7.1	La vérification informatique	23
7.2	Saisie des données	24
7.3	Nettoyage	24
7.4	Vérification de liens	25
7.5	Vérifications préliminaires	25
7.6	Vérification de cheminement	25
7.7	Codage des questions ouvertes	25
7.8	Vérification de la cohérence	26
7.9	Indicateurs d'imputation	26
7.10	Variables dérivées	27
7.11	Structure de codage des variables	27
7.12	Nomenclature	28
7.13	Siglaison des sections du questionnaire	29
7.14	Exemples de noms de variables	31

8.0	Contenu de l'enquête.....	33
8.1	Composantes de l'enquête	33
8.2	Variables démographiques	33
8.3	Questionnaire de l'adulte	34
8.4	Questionnaire de l'enfant	36
8.5	Échelles de l'enquête	37
8.5.1	Échelle de comportement	38
8.5.2	Échelle de dépression.....	39
8.5.3	Échelle de fonctionnement de la famille	39
8.5.4	Échelle de sécurité du quartier	40
8.5.5	Échelles du rôle parental	40
8.5.6	Échelle du soutien social	41
9.0	Qualité des données	43
9.1	Taux de réponse	43
9.2	Erreurs relatives à l'enquête	43
9.3	Base de sondage	43
9.4	Répondants hors du champ de l'enquête	44
10.0	Lignes directrices pour la totalisation, l'analyse et la diffusion de données.....	47
10.1	Lignes directrices pour la pondération de l'échantillon en vue de la totalisation	47
10.1.1	Définitions de types d'estimations : catégoriques et quantitatives	47
10.1.2	Totalisation d'estimations catégoriques	48
10.1.3	Totalisation d'estimations quantitatives	49
10.2	Lignes directrices pour la diffusion de coefficients de variation.....	49
11.0	Variabilité d'échantillonnage	51
12.0	Pondération	53

1.0 Introduction

Statistique Canada a mené l'Enquête sur les communautés entre février et juin 2005 avec l'appui et la collaboration de Ressources humaines et Développement social Canada (anciennement Développement social Canada et Développement des ressources humaines Canada). Ce manuel a été produit pour faciliter la manipulation du fichier de microdonnées portant sur les résultats de l'enquête.

Toutes les questions concernant l'ensemble de données ou son utilisation devraient être adressées à :

Statistique Canada

Services à la clientèle
Division des enquêtes spéciales
2500 — R, immeuble Principale
Ottawa (Ontario) K1A 0T6
Téléphone : (613) 951-3321 ou appelez sans frais : 1 800 461-9050
Télécopieur : (613) 951-4527
Courriel : des@statcan.ca

2.0 Contexte

Il est de plus en plus reconnu que les premières années d'un enfant (de 0 à 5 ans) sont d'une importance capitale pour son développement. Ces années façonnent les résultats à long terme liés aux réalisations scolaires, au succès en matière d'emploi, la santé, la qualité de vie et la capacité d'adaptation. C'est pour cette raison et pour donner suite à l'engagement du gouvernement du Canada à assurer que tous les enfants sont prêts à apprendre et à prendre part en société, Ressources humaines et Développement social Canada parraine Comprendre la petite enfance (CPE), un projet national de recherche visant à :

- enrichir nos connaissances relatives aux facteurs qui influencent sur le développement harmonieux de l'enfant,
- suivre les progrès de notre société dans l'amélioration des résultats des jeunes enfants, et
- catalyser les initiatives communautaires.

Le projet CPE, qui est une initiative conjointe entre Statistique Canada, Ressources humaines et Développement social Canada, et les communautés impliquées, fait partie de cette initiative. Ce projet aidera les communautés à améliorer les résultats des enfants et des familles qui y vivent en leur fournissant, à l'échelon communautaire, l'information nécessaire à l'élaboration et à l'amélioration de stratégies locales visant le développement harmonieux de l'enfant.

La population visée par le projet CPE est les enfants qui vivent dans les communautés allant à la maternelle (ou d'un type d'établissement d'enseignement équivalent) au moment de l'étude. Le projet se compose des trois composantes suivantes :

1. L'instrument de mesure du développement du jeune enfant

L'instrument de mesure du développement du jeune enfant (IMDJE) est un questionnaire qui a été développé et administré par le Canadian Centre for Studies of Children at Risk de l'université McMaster. Il mesure une variété de facteurs qui ont un impact sur la maturité pour l'apprentissage de l'enfant, entre autre le bien-être physique, les capacités linguistiques et cognitives, ainsi que le développement affectif et social. Ce questionnaire est rempli par les enseignants de la maternelle pour tous les enfants à la maternelle de la communauté impliqués dans ce projet. Les résultats de l'IMDJE ont été fusionnés avec les données de l'Enquête sur les communautés lorsque le parent l'a autorisé.

2. L'Enquête sur les communautés

Un échantillon d'enfants est sélectionné pour participer à l'Enquête sur les communautés parmi tous les enfants fréquentant la maternelle. La collecte des données se déroule ensuite en deux étapes : la première consiste à interviewer les parents des enfants par téléphone, tandis que la seconde consiste à soumettre les enfants à diverses mesures d'évaluation directes. Les bureaux régionaux de Statistique Canada sont responsables d'entreprendre la collecte de l'Enquête sur les communautés.

3. L'étude ethnographique

Cette composante consiste en une description de la communauté, et elle est réalisée directement par les coordonnateurs de recherche de la communauté. Cette partie de l'étude fait la description de diverses caractéristiques de la communauté, entre autres les services offerts, les parcs et aires de jeux et autres ressources communautaires.

Le projet CPE s'étend sur cinq ans. Les trois composantes – l'instrument de mesure du développement du jeune enfant, l'Enquête sur les communautés et l'étude ethnographique – sont complétées au cours de la première année, soit 2000. Au cours des 2^e, 3^e et 4^e années (2001 à 2003), seul l'instrument de

mesure du développement du jeune enfant est administré auprès des enfants de la maternelle (ou l'équivalent). Au cours de la 5^e année (2004), et de la 6^e année (2005) toutes les trois composantes sont de nouveau complétées.

En 1999, North York (Ontario) a été la première communauté à prendre part à cette initiative en tant que projet pilote. Elle a été suivie par cinq autres communautés au printemps 2000, soit Coquitlam (Colombie-Britannique), Prince Albert (Saskatchewan), Winnipeg (Manitoba), sud-ouest de Terre-Neuve, et l'Île-du-Prince-Édouard. Parmi les cinq communautés étudiées en 2000, seulement quatre ont été suivies en 2004 (Prince Albert, Winnipeg, sud-ouest de Terre-Neuve et l'Île-du-Prince-Édouard) et les résultats ont été rapportés en 2004. Les sept communautés qui ont été choisies participer à 2001 et au suivi dans 2005 sont : Abbotsford (Colombie-Britannique), Saskatoon (Saskatchewan); South Eastman (Manitoba), Mississauga (Ontario), Niagara Falls (Ontario), Montréal (Québec) et Hampton (Nouveau-Brunswick).

Ce rapport va noter les résultats des données d'Enquête sur les communautés de 2005. Noter que cette sixième enquête (2005) n'est pas administré au même groupe des enfants qu'en année deux (2001). Au lieu de cela, l'enquête de suivi est accomplie un nouveau groupe des enfants de la maternelle (ou équivalent) puisque c'est la communauté elle-même plutôt que les enfants dans la communauté qui est notre intérêt principal.

3.0 Objectifs

L'Enquête sur les communautés de 2001 avait pour but, d'une part, de déterminer les influences de la communauté et les autres facteurs pouvant avoir une incidence sur des résultats à long terme comme les réalisations scolaires, le succès en matière d'emploi, la santé, la qualité de vie et la capacité d'adaptation dans certaines communautés et, d'autre part, de recueillir de l'information s'y rattachant.

À ce moment-là, les quatre objectifs principaux étaient de :

- fournir des données exhaustives sur les déterminants du développement des enfants, y compris des renseignements détaillés sur le rôle de la famille, les caractéristiques de la communauté et les arrangements relatifs à la garde des enfants par d'autres personnes que les parents;
- acquérir une meilleure connaissance des facteurs qui influencent la maturité pour l'apprentissage de l'enfant à son entrée dans le système scolaire;
- comprendre les répercussions de l'influence de la communauté sur d'autres facteurs (p. ex. individu et famille) dans la vie de l'enfant;
- permettre aux communautés sélectionnées de comparer les caractéristiques de leurs enfants aux normes nationales établies par l'Enquête longitudinale nationale sur les enfants et les jeunes.

Les objectifs de l'Enquête sur les communautés de 2005 demeurent les mêmes que ceux de l'Enquête sur les communautés de 2001. En outre, les chercheurs ont la possibilité de mesurer tout changement ayant pu survenir relativement à la maturité scolaire des enfants dans chaque communauté en comparaison avec les données repères établies dans le cadre de l'Enquête sur les communautés de 2001.

4.0 Concepts et définitions

Ce chapitre donne un aperçu des concepts et des définitions d'intérêt pour les utilisateurs. Nombre de variables et de concepts sont essentiels aux analyses des données de l'Enquête sur les communautés de 2005. Une explication des concepts clés est présentée ci-dessous. Veuillez noter que les variables dérivées sont celles qui ne proviennent pas directement des renseignements fournis par les répondants, mais qui sont plutôt calculées à partir de ces renseignements.

4.1 Unité d'analyse

Pour chaque enfant de l'enquête, une foule de renseignements est recueillie sur la famille de l'enfant, sur le ou les parents et sur le quartier. L'enfant est considéré comme l'unité d'analyse et les variables de la famille et du ménage doivent faire partie des caractéristiques de l'enfant. Ainsi, les fichiers qui ont été constitués pour toutes les données de l'Enquête sur les communautés se composent d'enregistrements des enfants. Afin de comprendre la situation de la famille, les estimations tels que le nombre d'enfants dans les familles monoparentales peuvent être produites.

4.2 Personne la mieux renseignée et conjoint

Au moment de l'interview téléphonique, la personne-ressource du ménage (soit la personne dont le nom figurait dans la base de sondage) est demandée d'identifier quelle était la **personne la mieux renseignée (PMR)** sur l'enfant dans le ménage. Celle-ci devait fournir des renseignements sur l'enfant sélectionné au sein du ménage ainsi que des renseignements sociodémographiques sur elle-même et sur son/sa conjoint(e)/partenaire, si présent.

Voici la ventilation des liens entre les PMR et les enfants de l'Enquête sur les communautés de 2005¹ :

- Pour 87,9 % des enfants répondants, la PMR était la mère (la mère biologique dans 86,0 % et la mère par alliance, adoptive ou de famille d'accueil dans 1,9 % des cas).
- Pour 10,8 % des enfants, la PMR était le père (le père biologique dans 10,5 % et le père par alliance, adoptif ou de famille d'accueil dans 0,3 % des cas).
- Pour 1,3 % des enfants, la PMR n'était pas leur parent.

Dans les cas où la PMR n'était pas un des parents, l'enfant vivait le plus souvent dans le même ménage qu'un de ses parents, mais ce dernier n'avait pas été retenu comme PMR. Dans la plupart des cas, cette situation avait lieu lorsqu'il y avait des barrières de communication avec la PMR mais un autre membre du ménage pouvait parler la langue et fournir l'information nécessaire. Cette situation a lieu aussi quand la mère était très jeune et vivait elle-même avec ses parents, c'est-à-dire les grands-parents de l'enfant, et que la grand-mère avait été sélectionnée comme PMR.

Lorsque la PMR avait un partenaire qui résidait sous le même toit au moment de l'interview, cette personne était désignée comme le **conjoint**. Étaient considérés comme conjoints, les partenaires mariés et les partenaires en union libre. Des renseignements socio-économiques détaillés ont été recueillis au sujet du conjoint/partenaire afin de décrire la situation de la famille de l'enfant.

¹ Ces pourcentages estimés concernant la PMR et le conjoint/partenaire sont fondés sur des données pondérées provenant du fichier partage.

Voici la ventilation des liens entre les conjoints/partenaires et les enfants de l'Enquête sur les communautés de 2005 :

- Pour 22,4 % des enfants, la PMR n'avait pas de conjoint/partenaire vivant sous le même toit.
- Pour 67,8 % des enfants, le conjoint/partenaire était le père (64,4 %, le père biologique et 3,4 %, le père par alliance, adoptif ou de famille d'accueil).
- Pour 9,1 % des enfants, le conjoint/partenaire était la mère (8,8 %, la mère biologique et la mère par alliance, adoptive ou de famille d'accueil dans 0,3 % des cas).
- Pour les 0,7 % d'autres enfants, le conjoint/partenaire n'était pas leur parent.

4.3 Variables dérivées relatives à la famille

Il est possible de décrire la famille d'un enfant de plusieurs façons à partir des données de l'Enquête sur les communautés. Bon nombre des variables relatives à la famille utilisées pour décrire les enfants de l'Enquête sur les communautés ont été tirées de ce qu'il est convenu d'appeler la grille des liens. Dans le cadre du questionnaire du ménage, certaines données démographiques de base ont été recueillies pour tous les membres du ménage de l'enfant. Faisant partie intégrante de ce questionnaire, des questions étaient posées sur les liens de chaque personne avec chacun des autres membres du ménage. À partir de cette information, il a été possible de créer un ensemble exhaustif de variables en vue de décrire la situation familiale de l'enfant.

Certaines des variables dérivées relatives à la famille figurent ci-après. Les titres de ces variables sont entre parenthèses.

Famille monoparentale

Les enfants qui vivent avec un parent sont classés comme membres de familles monoparentales.

En Communautés 2005, 77,8 % des enfants vivaient avec deux parents, 21,6 % vivaient avec un seul parent, et 0,6 % ne vivaient avec aucun de leurs parents (GDMCD04).²

Famille intacte, famille reconstituée et famille recomposée

Les enfants qui vivent avec deux parents sont classés comme membres d'une famille intacte, reconstituée et (ou) recomposée selon les liens des ces enfants avec les parents³.

² Ces estimations pour les variables dérivées relatives à la famille sont basées sur des données pondérées provenant du fichier partage.

³ Les enfants en famille d'accueil et les enfants qui vivent avec un seul parent ne sont pas classés dans les familles reconstituées, recomposées ou intactes. Pour les familles recomposées, intactes ou reconstituées, lorsqu'un enfant était l'enfant adoptif d'un des parents et l'enfant biologique de l'autre parent, on le considérait comme un enfant de famille reconstituée, de sorte que sa famille se voyait attribuer le statut de famille reconstituée. Dans d'autres publications de Statistique Canada, les enfants qui se trouvent dans cette situation sont considérés comme les enfants biologiques des deux parents.

Famille intacte

Une famille intacte est formée d'un couple marié ou en union libre dont tous les enfants sont les enfants biologiques et (ou) adoptifs des deux membres du couple.

Dans le cadre de l'Enquête sur les communautés, 70,4 % des enfants vivaient dans des familles intactes (GDMCD16).

Famille reconstituée

Une famille reconstituée est formée d'un couple marié ou en union libre, avec au moins un enfant qui vit avec eux et est l'enfant biologique ou adoptif d'un seul des parents. Il convient de souligner que l'enfant biologique des deux parents est considéré comme faisant partie d'une famille reconstituée si au moins un de ses parents vit avec un enfant par alliance.

Dans le cadre de l'Enquête sur les communautés de 2005, 2,8 % des enfants étaient des enfants par alliance (GDMCD03), et 7,4 % vivaient dans des familles reconstituées (GDMCD15).

Famille recomposée

Une famille recomposée est formée d'un couple marié ou en union libre qui vit avec au moins deux enfants, dont l'un n'a pas les mêmes parents biologiques et (ou) adoptifs que l'autre ou les autres enfants. Voici des exemples de familles recomposées :

- un couple qui vit avec les enfants biologiques de la femme et avec les enfants biologiques de l'homme (c'est-à-dire les enfants de la femme et ceux de l'homme);
- un couple qui vit avec les enfants biologiques de la femme et avec les enfants issus de la nouvelle union (c'est-à-dire les enfants de la femme et ceux du couple).

Les familles recomposées forment un sous-ensemble des familles reconstituées. Dans le cadre de l'Enquête sur les communautés de 2005, 5,8 % des enfants étaient membres de familles recomposées (GDMCD14).

Famille économique

Une famille économique est définie comme l'ensemble des membres de la famille unis par le sang, le mariage, l'union libre ou l'adoption; les enfants en famille d'accueil sont considérés comme membres de la famille économique. Par exemple, si une femme vit dans un ménage avec son conjoint et deux enfants, en plus de sa sœur et de l'enfant de sa sœur, toutes ces personnes sont incluses dans une famille économique. S'il y avait également dans le ménage un pensionnaire et son enfant, ils formeraient une deuxième famille économique.

Fratricie

La fratrie englobe les frères et les sœurs germains, les demi-frères et demi-sœurs, ainsi que les frères et sœurs par alliance, adoptés ou en famille d'accueil. Seuls les frères et sœurs qui vivent dans le ménage sont pris en compte dans le calcul des variables dérivées relatives aux fratries. Dans le cas des unions libres, si les deux membres du couple ont avec eux leurs propres enfants, ces derniers sont considérés comme étant frères et sœurs. Tous les frères et sœurs vivant dans le ménage, y compris les frères et sœurs adultes, sont pris en compte dans le calcul des variables dérivées relatives aux fratries. Ces variables comprennent le nombre de frères et

sœurs plus âgés, les frères et sœurs plus jeunes ou les frères et sœurs ayant exactement la même date de naissance, par exemple les jumeaux (GDMCD08, GDMCD09, GDMCD10 et GDMCD11).

4.4 Variables dérivées relatives au statut socio-économique

Dans l'Enquête sur les communautés (2001) antérieure, deux variables dérivées avaient été calculées pour aider les analystes à comprendre et à expliquer la situation socio-économique de la famille de l'enfant : le statut socio-économique (SSE) et le ratio de revenu.

Statut socio-économique

Cette variable dérivée qui a été calculée pour l'Enquête sur les communautés de 2001, ne sera pas disponible pour l'Enquête sur les communautés de 2005. Le système de classification regroupe les professions décrites dans la Classification type des professions (CTP 1980) de 1980 de Statistique Canada en 16 catégories à peu près homogènes qui classe les professions selon leur statut social ou leur prestige. Depuis ce temps, cette structure de codage a été révisée (CTP 1991), mais par contre aucune nouvelle définition du statut socio-économique n'a été développée en utilisant cette nouvelle classification.

Ratio de revenu

On a créé deux variables dérivées (GINHD04A et GINHD05A) afin de comparer le revenu du ménage au seuil de faible revenu (SFR). Les seuils de faible revenu servent à distinguer les unités familiales qui sont à « faible revenu » des « autres » groupes. Une unité familiale est considérée à faible revenu lorsque son revenu est inférieur à la valeur du seuil correspondant à sa taille d'unité familiale et à la taille de sa communauté. Toute famille ayant un revenu égal ou supérieur au seuil de faible revenu est considérée dans la catégorie « autre »⁴. La variable GINHD03A donne la valeur du SFR par zone géographique.

4.5 Imputation du revenu

On observe de nombreux cas de non-réponse partielle chez les répondants. Il peut s'agir d'une non-réponse à une composante dans son ensemble ou à certaines questions seulement. L'imputation est la procédure qui consiste à remplacer les valeurs manquantes ou incohérentes par des valeurs acceptables. Dans le cadre de l'Enquête sur les communautés de 2005, on effectue une procédure d'imputation pour certaines variables dans la section relative au revenu des adultes

On a intégré des indicateurs d'imputation au fichier de l'Enquête sur les communautés de 2005 pour informer les utilisateurs de l'ampleur de l'imputation et des items particuliers ayant fait l'objet d'une imputation dans tel ou tel enregistrement. Pour tous les indicateurs d'imputation du fichier de données des communautés, un « I » apparaît au cinquième ou sixième caractère du nom de la variable. Ainsi, GINHEI03 représente l'indicateur d'imputation pour le revenu du ménage (GINHEQ03).

Imputation du revenu des adultes

L'interview du ménage dans le cadre de l'Enquête sur les communautés de 2005 comporte plusieurs questions relatives au revenu. On recueille des renseignements sur le revenu de la PMR et de son conjoint selon les trois sources suivantes : salaires et traitements, revenu net de travail autonome, et prestations d'assurance emploi. On recueille aussi des renseignements sur le revenu du ménage selon les quatre sources suivantes : prestation fiscale pour enfants/prestation nationale pour enfants, assistance sociale, allocation de soutien des enfants ou au conjoint et autres sources.

⁴ Pour de plus amples renseignements sur les mesures utilisées par Statistique Canada en matière de faible revenu, veuillez consulter *Seuils de faible revenu de 1994 à 2003* et *Mesures du faible revenu de 1992 à 2001*, N° 75F0002MIF, 2004002 au catalogue.

Le revenu est un sujet délicat. Certains répondants refusent donc de donner des réponses aux questions détaillées se rapportant au revenu. Parmi ceux-ci, certains fournissent une estimation du revenu total de leur ménage ou une estimation de leur revenu au moyen des intervalles de revenu. Enfin, parmi ceux qui répondent aux questions, il arrive que les montants indiqués aux sections touchant le revenu soient incompatibles avec les réponses données à la section relative à l'activité sur le marché du travail (par exemple, un répondant qui, selon les réponses fournies à la section relative à l'activité sur le marché du travail, a travaillé au cours des 12 derniers mois mais qui ne déclare pas de salaire ou de revenu net d'un emploi autonome dans la section sur le revenu). On procède à une imputation du revenu pour combler les valeurs manquantes attribuables à la non-réponse partielle ou pour corriger les données incohérentes lorsque cela est possible. L'imputation a été aussi faite pour les ménages qui ont déclaré un revenu total de moins de 6 000 \$.

L'imputation a été faite pour chacune des sources de revenu. L'imputation a été effectuée à l'aide de la méthode du plus proche voisin. Cette méthode consiste à retracer un répondant ayant fourni une réponse à la section sur le revenu (un donneur) et dont les caractéristiques sont semblables à celles de la personne ou le ménage n'ayant pas fourni de renseignements complets sur le revenu (un receveur). Une fois qu'on a identifié le voisin le plus proche, les montants déclarés par le donneur servent à imputer les sources de revenu manquantes du receveur. Deux types d'imputation ont été faits. Tout d'abord, les trois sources de revenu du PMR et du conjoint ont été imputées. Les autres variables ont été imputées au niveau du ménage.

L'imputation au niveau du ménage a été faite selon l'une des trois méthodes suivantes. Pour les ménages qui ont fourni une estimation du revenu du ménage, cette information a été utilisée pour trouver un donneur. Pour les ménages qui ont fourni une estimation du revenu du ménage à l'aide des intervalles de revenu, la valeur de l'intervalle a été utilisée pour trouver le donneur. Si aucune information n'était disponible au sujet du revenu, d'autres variables au niveau du ménage, telle que la province, ont été utilisées. Les indicateurs d'imputation indiquent quelle méthode d'imputation a été utilisée.

Indicateur d'imputation = 1	L'estimation du revenu total a été utilisée pour déterminer le donneur
Indicateur d'imputation = 2	Intervalle de revenu total a été utilisé pour déterminer le donneur
Indicateur d'imputation = 3	Aucune information additionnelle sur le revenu a été utilisée pour déterminer le donneur

Les taux d'imputation pour les variables du revenu se trouvent dans le tableau suivant.

Variable	Taux d'imputation
GINPc1AA Revenu de PMR provenant de salaires et traitements	22,7 %
GINPc1AB Revenu de PMR provenant d'un emploi autonome	22,3 %
GINPc1AC Revenu de PMR provenant des prestations d'assurance emploi	21,0 %
GINPCD04 Revenu total du PMR	26,5 %
GINSc1AA Revenu du conjoint provenant de salaires et traitements	26,5 %*
GINSc1AB Revenu de conjoint provenant d'un emploi autonome	23,8 %*
GINSc1AC Revenu du conjoint provenant des prestations d'assurance emploi	21,5 %*
GINSCD04 Revenu total du conjoint	29,6 %*

Variable	Taux d'imputation
GINHe1AD Revenu du ménage provenant des prestations fiscales pour enfants	26,0 %
GINHe1AE Revenu du ménage provenant de l'assistance sociale	21,0 %
GINHe1AF Revenu du ménage provenant de l'allocation de soutien des enfants ou au conjoint(e)	20,8 %
GINHe1AG Revenu du ménage provenant des autres sources	23,3 %
GINHEQ03 Revenu total du ménage	37,2 %

* Les ménages où il n'y avait pas un conjoint ne sont pas inclus dans le calcul du taux d'imputation.

4.6 Instrument de mesure du développement du jeune enfant

L'instrument de mesure du développement du jeune enfant (IMDJE) est un questionnaire qui a été développé et administré par le Canadian Centre for Studies of Children at Risk de l'université McMaster. Il s'agit d'une liste de vérification remplie par les enseignants et mesurant une variété de facteurs qui ont un impact sur la maturité pour l'apprentissage à l'école de l'enfant. Le terme « maturité pour l'apprentissage à l'école » renvoie à la capacité de l'enfant de respecter les règles de l'école, telles qu'être coopératif, rester assis sagement et écouter l'enseignant, et de profiter des activités éducatives offertes à l'école. Pour répondre au questionnaire, les enseignants ont recours aux observations qu'ils font au cours de plusieurs mois d'interaction avec l'enfant en classe. Même s'il est rempli sur une base individuelle, il ne doit être interprété qu'au niveau du groupe.

Les résultats du questionnaire sont saisis et cotés à l'université McMaster, qui les envoie ensuite à Statistique Canada. Lorsque les parents l'autorisent, les résultats de l'IMDJE sont couplés aux données de l'Enquête sur les communautés de 2005 au moyen d'un identificateur commun. Chacun des cinq domaines de maturité pour l'apprentissage est inclus sous forme d'échelle dans le fichier de données. Les domaines sont les suivants :

1. santé physique et bien-être (GEICdS4A),
2. connaissance et compétence sociales (GEICdS5A),
3. maturité et santé émotionnelle (GEICdS1A),
4. langue et développement cognitif (GEICdS3A), et
5. aptitudes de communication et connaissances générales (GEICdS2A).

5.0 Méthodologie de l'enquête

L'Enquête sur les communautés a été menée entre février et juin 2005. En février, mars et avril, les intervieweurs de Statistique Canada ont procédé à la collecte des données sur les ménages en effectuant des interviews téléphoniques auprès des parents des enfants. En mai et en juin, les intervieweurs de Statistique Canada se sont rendus dans les écoles pour procéder à l'étape de mesure directe auprès des enfants dont les parents avaient donné leur consentement.

5.1 Population visée

La population cible de chaque communauté comprend tous les enfants inscrits dans les écoles maternelles des commissions scolaires figurant dans le tableau ci-dessous à l'automne 2004 et qui allaient toujours à l'école de la communauté à l'hiver 2005 (au cours de la période de collecte des ménages).

Communautés	Associations ou commissions scolaires
South Eastman (Manitoba)	Border Land School Division Division scolaire franco-manitobaine Hanover School Division Red River Valley School Division Seine River School Division
Hampton (Nouveau-Brunswick)	School District # 6
Abbotsford (Colombie-Britannique)	School District # 34
Mississauga (Ontario)	Dufferin-Pell Catholic District School Board Peel District School Board
Niagara Falls (Ontario)	District School Board of Niagara Niagara Catholic District School Board
Saskatoon (Saskatchewan)	Saskatoon Catholic Schools Saskatoon Public Schools
Montréal (Québec)	Commission scolaire de Montréal

5.2 Plan d'échantillonnage

L'unité d'échantillonnage de l'Enquête sur les communautés est l'enfant. Les enfants sont sélectionnés parmi les dossiers scolaires fournis à Statistique Canada par les commissions scolaires de la communauté.

Lors de la sélection de l'échantillon des sept communautés participantes, Statistique Canada a obtenu une base de 9 807 enfants potentiels parmi lesquels choisir. Une vérification de la qualité de la base a été réalisée avant la sélection de l'échantillon et consistait à éliminer les enregistrements en double et à rectifier les renseignements absents, incomplets ou contradictoires. Un échantillon systématique d'enfants a été sélectionné aléatoirement à partir des listes, qui étaient classées par école et code postal, puis des renseignements ont été recueillis au sujet des personnes échantillonnées en préparation des interviews. Dans trois des sept communautés participantes, tous les enfants ont été sélectionnés pour les interviews, étant donné que la population d'enfants dans ces communautés était inférieure à 700 ou avoisinait ce nombre. L'échantillon prélevé initialement comprenait un certain pourcentage d'unités hors du champ d'observation (par exemple, des enfants qui étaient inscrits à la maternelle dans la communauté à l'automne 2004, mais qui ne l'étaient plus au moment de la période d'interview de l'Enquête sur les communautés). Il convient aussi de noter que certains enfants qui allaient à

l'école dans la communauté au moment de la période d'interview auprès des ménages mais qui n'étaient pas inscrits à l'école à l'automne 2004 ne pouvaient pas être inclus dans l'échantillon et ne sont pas pris en compte dans l'estimation.

5.3 Taille de l'échantillon

Le tableau ci-après donne la ventilation de la taille de la population et de l'échantillon selon chaque communauté.

Communautés	Population	Taille de l'échantillon
South Eastman (Manitoba)	809	582
Hampton (Nouveau-Brunswick)	697	697
Abbotsford (Colombie-Britannique)	1 489	582
Mississauga (Ontario)	668	668
Niagara Falls (Ontario)	721	721
Saskatoon (Saskatchewan)	2 307	584
Montréal (Québec)	3 116	575
Total	9 807	4 409

6.0 Collecte des données

La collecte des données dans les communautés étudiées s'est déroulée entre février et juin 2005. La collecte des données auprès des ménages a été effectuée entre février et avril par les intervieweurs de Statistique Canada, à l'aide de l'application de l'interview téléphonique assistée par ordinateur (ITAO), qui ont communiqué avec les parents de l'enfant et ont mené les interviews par téléphone. En mai et juin, les intervieweurs sont allés dans les écoles afin de faire passer la partie de l'enquête comportant les mesures directes aux enfants dont les parents avaient donné leur consentement. L'instrument de mesure du développement du jeune enfant (IMDJE) a été rempli, à l'école, par les enseignants entre janvier et mars.

6.1 Processus relatif au consentement éclairé

En septembre 2004, les commissions scolaires de chaque communauté ont reçu une lettre décrivant le projet Comprendre la petite enfance (CPE). La lettre expliquait la procédure qui serait adoptée et demandait aux commissions scolaires de collaborer en fournissant à Statistique Canada une liste de noms des classes de maternelle lui permettant de créer une base de sondage dans laquelle sélectionner un échantillon d'enfants.

Des lettres de présentation ont également été envoyées aux enseignants et aux directeurs des écoles concernées pour demander leur collaboration, de même qu'aux parents dont les enfants fréquentaient la maternelle relevant de la commission scolaire. On informait les parents que leur enfant pourrait être sélectionné pour participer à l'Enquête sur les communautés et on les priait de signer un formulaire de consentement permettant à leur enfant de participer aux tests de mesures directes s'il était sélectionné.

Statistique Canada a ensuite choisi un échantillon d'enfants au hasard dans la liste de noms des classes de maternelle. Une lettre de suivi a alors été envoyée aux parents des enfants sélectionnés, précisant le moment où ils pourraient s'attendre à recevoir l'appel d'un intervieweur. Au cours des interviews téléphoniques, qui ont commencé au début février 2005, les parents étaient priés de confirmer de vive voix qu'ils autorisaient leur enfant à participer aux tests de mesures directes. Les intervieweurs de Statistique Canada leur demandaient également l'autorisation de communiquer les données de l'enquête à Ressources humaines et Développement social Canada et de coupler les résultats de l'enquête aux résultats de l'IMDJE.

En mai et en juin, les intervieweurs de Statistique Canada se sont rendus dans les écoles afin de soumettre les enfants dont les parents avaient donné leur consentement, de vive voix ou par écrit, aux tests de mesures directes. Les enseignants de maternelle ont rempli le questionnaire de l'IMDJE pour chaque élève de leur classe (à l'exception de ceux dont les parents avaient refusé leur consentement) entre les mois de janvier et de mars.

6.2 Collecte des données auprès des ménages

Dans le cas de la collecte des données auprès des ménages, les intervieweurs de Statistique Canada ont recueilli les données à l'aide de l'application ITAO auprès de la personne qui se déclarait être la personne la mieux renseignée (PMR) sur l'enfant, habituellement la mère. Une brève description de chaque section de l'interview assistée par ordinateur se trouve ci-dessous.

6.2.1 Composante « Entrée/Sortie »

Le premier contact pour ces ménages a été établi grâce à l'adresse et au numéro de téléphone trouvés dans les fichiers fournis par les commissions scolaires. Après avoir confirmé la présence de l'enfant sélectionné dans le ménage, l'intervieweur poursuivait l'interview en dressant une liste de tous les membres du ménage, en recueillant des

renseignements de dépistage et en obtenant les caractéristiques démographiques de base, notamment le sexe, la date de naissance, l'état matrimonial, les liens entre les membres du ménage et les renseignements sur le logement.

6.2.2 La composante « Enfant »

Le questionnaire de l'enfant était rempli pour l'enfant en maternelle sélectionné dans le ménage. La PMR répondait aux questions pour l'enfant sur les sujets suivants :

- Santé
- Comportement
- Activités
- Alphabétisation
- Rôle parental
- Antécédents familiaux
- Garde des enfants
- Habitudes de sommeil
- Comportement positif
- Communication
- Utilisation des ressources de la communauté

6.2.3 La composante « Adulte » (Parent)

Les questions de cette section portaient sur la PMR et son/sa conjoint(e)/partenaire (le cas échéant). Elles abordaient les sujets suivants :

- Scolarité
- Revenu
- Population active
- Santé
- Fonctionnement de la famille
- Sécurité du quartier
- Caractéristiques sociodémographiques
- Soutien social

6.3 Mesures cognitives recueillies auprès de l'enfant

Trois tests de mesures directes ont été administrés à chacun des enfants échantillonnés par un intervieweur de Statistique Canada dans les écoles : l'Échelle de vocabulaire en images de Peabody (EVIP), ainsi que des versions abrégées du Qui suis-je? et de la Connaissance des nombres. Les tests prenaient environ 30 minutes en tout, et seuls les enfants dont les parents avaient donné leur consentement y étaient soumis.

Les enfants incapables de communiquer en anglais ou en français ne participaient pas aux tests.

Échelle de vocabulaire en images de Peabody

L'EVIP-R a servi à mesurer le niveau de vocabulaire passif, ou vocabulaire compris, de l'enfant. Tous les enfants sélectionnés y ont participé. Les enfants devaient indiquer du doigt l'image qu'il croyait (parmi un choix de quatre) correspondre au mot du test que l'intervieweur avait prononcé. La difficulté des images et des mots augmentait à mesure que le test progressait. Les résultats sont prédicteurs des succès scolaires de l'enfant. Le test a été conçu par Lloyd et Leota Dunn de l'université de Hawaii et est largement utilisé

pour évaluer le vocabulaire passif de tous les groupes d'âge (de 2,5 ans à l'âge adulte). Pour de plus amples renseignements visant la méthode de calculer les scores du EVIP-R, veuillez faire référence au chapitre 16.0 du Guide de l'utilisateur des microdonnées de l'Enquête longitudinale nationale sur les enfants et les jeunes, cycle 5.

Qui suis-je?

Le questionnaire Qui suis-je évalue l'enfant au moment où il entre à l'école, plus précisément les processus cognitifs qui sous-tendent l'apprentissage initial de la lecture, de l'écriture et du calcul. Le test original comprend trois échelles : la reproduction des formes géométriques (cercle, croix, carré, triangle et losange), la connaissance des symboles (écrire son nom, des lettres, des chiffres, des mots et des phrases) et le dessin (autoportrait). Cependant, à cause des contraintes de temps, le dessin a été retiré de l'Enquête sur les communautés. L'enfant doit exécuter le plus grand nombre de tâches possible dans le livret utilisé par l'évaluateur qui tourne les pages et donne les directives. L'instrument a été conçu par D^r Molly de Lemos et ses collègues de l'Australian Council for Educational Research et peut servir à évaluer les enfants de 3 à 7 ans.

Connaissance des nombres

Ce test évalue la compréhension qu'a l'enfant du concept des quantités et du système des nombres entiers. En général, l'enfant qui possède ces connaissances intuitives en commençant l'école réussit bien en mathématiques. L'enfant qui n'a pas cette compréhension ou qui travaille dans une langue autre que sa langue maternelle a souvent du mal à maîtriser les bases de l'arithmétique et à démontrer le sens des nombres. Ce test a été créé par D^r Robbie Case de l'Institut d'études pédagogiques de l'Ontario de l'université de Toronto. Il peut être administré à des enfants de 3,5 à 10,5 ans.

D^r Case et son confrère D^r Yukari Okamoto de l'université de la Californie ont élaboré une version abrégée de ce test pour l'Enquête longitudinale nationale sur les enfants et les jeunes. Ce test est administré oralement et les questions sont posées jusqu'à ce que l'enfant soit incapable de résoudre plus de la moitié des problèmes du niveau.

6.4 Durée de l'interview

Dans le cas de la collecte des données auprès des ménages, la durée de l'interview des ménages répondants de l'Enquête sur les communautés était d'environ une heure quarante minutes. L'administration des questionnaires de l'enfant et de l'adulte a duré en moyenne 70 minutes à compléter. L'intervieweur avait besoin de temps supplémentaire pour présenter l'enquête, remplir la liste des membres du ménage, etc.

Les tests cognitifs menés dans les écoles ont duré environ 10 minutes chacun, pour un total de 30 minutes par enfant.

6.5 Formation, supervision et contrôle relativement à l'interview

L'Enquête sur les communautés de 2005 a été effectuée par des intervieweurs de l'Enquête sur la population active (EPA). Tous les intervieweurs de l'EPA travaillent sous la supervision d'un groupe d'intervieweurs principaux. Ceux-ci ont pour responsabilité de s'assurer que les intervieweurs connaissent bien les concepts et les procédures de l'enquête. Ils ont aussi pour fonction de contrôler régulièrement le travail des intervieweurs et d'examiner les documents complétés par ces derniers. Les intervieweurs principaux s'assurent que des mesures de suivi sont prises rapidement dans les cas de refus et de non-réponse. Au besoin, les cas de non-

réponse sont confiés à l'intervieweur principal et réaffectés. Les intervieweurs principaux sont, à leur tour, supervisés par les gestionnaires du programme de l'EPA, qui se trouvent dans des bureaux régionaux de Statistique Canada.

Dans le cas de l'Enquête sur les communautés de 2005, une combinaison de la formation en salle de classe et de matériel autodidactique assurait que les intervieweurs comprenaient bien les concepts de l'enquête. Dans le cadre de la portion de l'auto-apprentissage, les intervieweurs lisaient le Manuel de l'intervieweur et complétaient des exercices à domicile avant la séance de formation d'une journée. La formation en classe sur la portion de l'enquête auprès des ménages a été dispensée aux bureaux régionaux par des employés du bureau central. Au cours de cette formation, un gestionnaire de programme ou un intervieweur principal a présenté un aperçu de l'enquête, a procédé à une enquête fictive auprès des participants, et a présenté des exercices en vue de permettre aux intervieweurs de réduire les cas de non-réponse.

Un intervieweur principal a donné une séance de formation d'une journée dans les écoles sur le volet des mesures directes de l'enquête. Dans ce cas encore, les intervieweurs ont dû lire leur manuel pour se familiariser avec le matériel avant la formation. Chaque participant a reçu une vidéocassette lui permettant de prévisionner le déroulement des mesures directes.

7.0 Traitement des données

Le principal produit de l'Enquête sur les communautés de 2005 est un fichier de micros données « épuré ». Ce chapitre présente un bref résumé des phases de traitement inhérentes à la production de ce fichier.

L'Enquête sur les communautés de 2005 utilise essentiellement des interviews téléphoniques assistées par ordinateur qui présentent plusieurs avantages lorsque débute le traitement des données, dont un des avantages mais non le moindre est de réduire la saisie de données. La vérification de l'enchaînement des données et la vérification de la cohérence sont programmées en direct dans l'application, ce qui veut dire que les données sont disponibles plus rapidement pour le traitement au bureau central et que la quantité de traitement est réduite.

Lorsque le fichier de données est reçu au bureau central, le traitement des données de l'Enquête sur les communautés de 2005 est effectué à l'aide d'un environnement de traitement généralisé. Il s'agit d'un système générique qui suit une série d'étapes pour « nettoyer » un fichier du début à la fin. Les étapes principales sont :

- Nettoyage
- Vérifications de liens
- Vérifications préliminaires
- Vérifications de l'enchaînement des questions
- Codage
- Vérifications de la cohérence
- Variables dérivées
- Fichier de traitement final

7.1 La vérification informatique

Tel qu'indiqué ci-haut, tous les renseignements auprès des ménages ont été obtenus au moyen d'une interview téléphonique à l'aide d'une application d'interviews assistées par ordinateur (IAO). On a pu ainsi inclure diverses fonctions de vérification dans le questionnaire pour les diverses composantes IAO des ménages de sorte que les renseignements recueillis soient de grande qualité. Voici quelques exemples illustrant le genre de vérifications faites dans le cadre du processus d'interviews assistées par ordinateur:

Écrans de révision

On a prévu des écrans de révision pour les données importantes et complexes. Par exemple, la procédure de sélection de la personne la mieux renseignée (PMR), élément clé de l'enquête, s'appuyait sur la liste des membres du ménage. La page-écran correspondante présentait des données démographiques sur chaque membre et son lien avec chacun des autres membres du ménage. Les données recueillies s'affichent à l'écran pour que l'intervieweur les confirme auprès du répondant avant de continuer l'interview.

Vérifications d'intervalles

On a également intégré des vérifications d'intervalles dans le système IAO pour les questions exigeant d'indiquer des valeurs numériques. Si les chiffres indiqués ne s'inscrivaient pas dans l'intervalle, une fenêtre apparaît instantanément pour indiquer l'erreur et demander à l'intervieweur de corriger la réponse erronée. Par exemple, si à la question portant sur le poids de l'enfant à la naissance, la valeur indiquée est trop élevée ou trop basse, un message s'affiche et demande à l'intervieweur de vérifier la réponse auprès du répondant.

Vérifications de cheminement de questions

Tous les cheminements de questions sont en intégration automatique au système IAO. Par exemple, à la section sur la garde de l'enfant, on demande à la PMR si elle (ou son conjoint/partenaire) place l'enfant en garderie ou en gardiennage afin de pouvoir travailler ou étudier. Le cheminement des questions peut varier selon la réponse. S'il y a utilisation de garderie ou gardiennage, le système IAO pose une suite de questions sur les modes de garde particuliers qui sont utilisés pour l'enfant. Dans le cas contraire, il saute automatiquement ces questions.

Vérifications de cohérence générale

On a prévu un certain nombre de vérifications de cohérence dans le système IAO, et les intervieweurs étaient en mesure de revenir à des questions déjà posées pour rectifier les incohérences. Les intervieweurs recevaient aussi des instructions à l'écran pour traiter ou régler des problèmes de réponse incomplète ou erronée. Par exemple, dans la collecte des données de la section sur la population active, le nombre de semaines passées à travailler, à ne pas travailler et à chercher du travail ne doit pas être de plus de 53 semaines. En cas de dépassement, le système signalait l'erreur et demandait à l'intervieweur de revenir à la question pour vérifier les données et apporter les corrections nécessaires.

Vérifications de cohérence

On a également effectué des vérifications de cohérence. Si des incohérences étaient relevées, le système demandait à l'intervieweur de vérifier les données auprès du répondant par une suite de questions. Par exemple, dans la section Socio-démographiques, si le répondant avait déclaré Canada comme le pays de naissance, le programme de vérifications de cohérence va automatiquement choisir la réponse « Canada, citoyen par naissance » lorsque demandé « Vous êtes citoyen de quel pays? ».

7.2 Saisie des données

Questionnaires papier

Certains questionnaires n'ont pas été remplis en utilisant des interviews assistées par ordinateur. Par exemple, le questionnaire Qui suis-je? qui fait partie des mesures directes a été rempli par l'enfant à l'aide d'un questionnaire papier.

L'instrument de mesure du développement du jeune enfant a été administré par l'Université McMaster, qui a traité les résultats et fourni le fichier de données à Ressources humaines et Développement social Canada, qui l'a ensuite transmis à Statistique Canada pour couplage au fichier de l'Enquête sur les communautés.

Saisie de données relative aux questionnaires papier

Les questionnaires Qui suis-je? ont été notés au bureau central de Statistique Canada. Les résultats de la notation ont été fournis aux préposés au traitement du bureau central en vue d'être traités plus en profondeur. Des vérifications de la qualité ont été intégrées au système de saisie de façon à repérer les entrées inhabituelles.

7.3 Nettoyage

Exigences définitives

L'objet de cette étape est de retirer les enregistrements en doubles complets et de séparer les enregistrements ayant des numéros d'identification en double pour examen. Ensuite, les données sont divisées en réponse et non-réponse selon les critères prédéterminés.

On a examiné les questionnaires des répondants et des non répondants et on a établi des spécifications fondées sur cette analyse pour déterminer les enregistrements qui seraient éliminés en raison de la non-réponse. Essentiellement, s'il manquait des données essentielles ou s'il n'y avait pas de réponse à plus de la moitié des questions sur un enregistrement, il était éliminé du fichier.

À la fin de cette étape, les enregistrements sont traités par type de questionnaire. C'est-à-dire : le questionnaire de l'adulte, le questionnaire de l'enfant, et le questionnaire du ménage.

Variables manquantes

Pour toutes les variables manquantes des ménages, on indiquait « Non déclaré ». Si on n'avait pas assez de renseignements, le ménage était retiré de l'échantillon des répondants et on en faisait un cas de non-réponse.

7.4 Vérification de liens

La vérification de liens établit les liens entre les membres du ménage et crée les variables dérivées familiales. Cette étape exécute une série de vérifications standards contre les données des liens déclarées pour tous les membres d'un ménage en particulier; certaines incohérences sont réglées automatiquement par une application utilisant un ensemble de règles, tandis que d'autres incohérences sont identifiées pour une révision manuelle et pour un recodage. Une série de variables dérivées connexes sont produites à partir des vérifications de liens.

7.5 Vérifications préliminaires

Le but de la vérification préliminaire est de procéder à un agencement des données et d'effectuer une première révision. Pour tous les enregistrements où il manquait des valeurs (blancs) provenant de la collecte, on a inséré la valeur « 9 », « 99 », « 999 » pour indiquer qu'aucune information n'avait été recueillie. Les entrées « Ne sait pas » retournées par l'application IAO comme « 9 » sont modifiées à « 7 » à l'étape de la vérification préliminaire. De même, les questions ayant l'instruction « Inscrivez toutes les réponses qui s'appliquent » ont été modifiées pour transformer leurs valeurs en réponses Oui (1) ou Non (2). Enfin, toutes les réponses de texte ont été enlevées du fichier de traitement et mises de côté pour être traitées séparément.

7.6 Vérification de cheminement

Cette vérification consiste à reproduire le cheminement des questions du questionnaire. Les variables qui sont sautées en suivant l'enchaînement des questions sont changées de « Non déclaré » à « Enchaînement valide » (6, 96, 996...).

Lorsque l'instruction « passez à » suit une question à laquelle on a répondu, toutes les questions sautées sont indiquées comme « Enchaînement valide ». Lorsque l'instruction « passez à » suit une réponse « Ne sait pas » ou « Refus », toutes les questions sautées sont indiquées comme « Non déclaré ».

7.7 Codage des questions ouvertes

Les intervieweurs ont enregistré sur le questionnaire quelques éléments d'information sous forme de réponses à des questions ouvertes. Ainsi, à la section sur la population active, on posait à une PMR qui avait travaillé au cours des 12 derniers mois une suite de questions ouvertes sur l'emploi actuel ou l'emploi le plus récent :

- Quel genre d'entreprise, de service ou d'industrie s'agit-il/s'agissait-il?
- Quel genre de travail faites-vous/ faisiez-vous?

- À cet emploi, quelles sont/étaient vos fonctions ou activités les plus importantes?

Mode d'enregistrement

L'intervieweur inscrivait littéralement les réponses données par le répondant. Au bureau central, les énoncés écrits ont été convertis en codes d'industrie et de profession pour décrire la nature du travail du répondant. On a recueilli des données semblables pour le conjoint/partenaire et attribué des codes descriptifs pour le genre de travail.

Mode de codage

Les questions ouvertes ont été codées au moyen de plusieurs classifications types. Les questions portant sur les professions ont été codées à l'aide de la Classification type des professions de 1991 (CTP) tandis que celles portant sur les industries l'ont été à l'aide du Système de classification des industries de l'Amérique du Nord (SCIAN) de 1997. On peut consulter les listes groupées de ces codes au fichier de données (GLFPcD7A et GLFPcD8A pour la PMR, et GLFSCD7A et GLFSCD8A pour le conjoint/partenaire).

7.8 Vérification de la cohérence

Après la vérification de l'enchaînement, nous avons procédé à une vérification de cohérence qui consiste à vérifier les liens entre deux variables ou plus. On utilise des tables de décision pour préciser les vérifications. Le logiciel LogiPlus est utilisé pour générer le code SAS en utilisant l'information fournie par les tables de décision. Un rapport comprenant les comptes « Avant » et « Après » des variables est produit, de même qu'un rapport comprenant le nombre de décisions prises.

Par exemple, à la section Socio-démographiques, une question portait sur la date de la première immigration au Canada (GSDCQ3) dans le cas des enfants nés à l'étranger. Une vérification de cohérence a permis de comparer cette réponse à l'année de naissance de l'enfant. Si l'année d'immigration précédait l'année de naissance, on indiquait « Non déclaré » pour l'année d'immigration dans le cadre de la vérification.

7.9 Indicateurs d'imputation

Pour diverses raisons, certaines variables sont absentes pour des ménages répondants compris dans le fichier de l'Enquête sur les communautés de 2005. On désigne ceux-ci généralement comme étant la non-réponse à une question ou la non-réponse partielle.

Pour certaines variables du fichier, on a procédé à une imputation au lieu d'attribuer un code spécial de non-réponse. L'imputation consiste à déterminer des valeurs vraisemblables pour des réponses manquantes ou incohérentes. Pour l'Enquête sur les communautés de 2005, l'imputation a porté sur le revenu du ménage et sur celui de la PMR. Pour de plus amples détails sur l'imputation, consulter la section 4.5.

On a ajouté des indicateurs d'imputation au fichier pour que l'utilisateur connaisse l'ampleur de cette imputation ainsi que les items spécifiques qui ont été imputés sur un enregistrement particulier.

Tous les indicateurs d'imputation du fichier de microdonnées comportent un « I » en cinquième ou sixième position dans le nom de la variable. Par exemple, l'indicateur d'imputation du revenu de la PMR est GINPEI03.

Veillez noter que si une ou quelques-unes des variables d'une composante qui a été répondue présentaient de la non-réponse, cette ou ces variables n'ont pas été imputées, à l'exception des variables sur le revenu, telles que mentionnées ci-haut.

7.10 Variables dérivées

Combinaison de questions

Un certain nombre de variables ont été dérivées en combinant des questions afin de faciliter l'analyse des données. Par exemple, dans la section sur la garde des enfants, on pose à la PMR une série de questions sur les modes de garde utilisés pour l'enfant, afin de permettre à cette personne ou son conjoint/partenaire de travailler ou d'étudier. Pour chaque mode de garde, une question est posée sur le nombre d'heures par semaine passées par l'enfant à cet endroit. À partir de cette information, une variable est créée pour indiquer le principal mode de garde utilisé pour permettre à la PMR ou son conjoint/partenaire de travailler ou d'étudier. Elle est dérivée à partir du nombre d'heures de chaque mode de garde et définie en fonction du plus grand nombre d'heures.

Nom des variables dérivées

Toutes les variables dérivées dans le fichier de microdonnées reçoivent un « D » en cinquième position dans leur nom. Par exemple, le nom de la variable dérivée portant sur le « Nombre d'heures par semaines passées au principal mode de garde » est GCRCD02.

7.11 Normes pour la structure de codage

On a conçu des normes pour la structure de codage des variables afin de décrire certaines situations en toute cohérence pour toutes les variables. Voici ces diverses situations avec leurs codes descriptifs respectifs.

Refus

Au cours d'une interview assistée par ordinateur, le répondant peut choisir de refuser de répondre à une question particulière. Le système IAO comporte une touche de fonction sur laquelle l'intervieweur appuie pour indiquer un refus. L'information est enregistrée pour la question et transmise au bureau central.

Sur le fichier de données, une question refusée comporte un code « 8 ». Dans le cas d'une variable à un chiffre, le code est « 8 », pour une variable à deux chiffres « 98 », pour une variable à trois chiffres « 998 », etc.

Ne sait pas

Il se peut que le répondant ne connaisse pas la réponse à une question particulière. Là encore, le système IAO comporte une touche de fonction sur laquelle l'intervieweur appuie pour expliquer cette situation.

Dans le fichier de données, le code utilisé pour indiquer que le répondant ne connaissait pas la réponse à une question particulière est « 7 ». Dans le cas d'une variable à un chiffre, le code est « 7 », pour une variable à deux chiffres « 97 », pour une variable à trois chiffres « 997 », etc.

Sans objet

Dans certains cas, une question ne s'applique pas au répondant. Dans le fichier de données, on se sert des codes « 6 », « 96 », « 996 », etc., pour indiquer qu'une question ou une variable dérivée est sans objet.

Dans certains cas, une question ou une série de questions était sans objet. Par exemple, la question sur le nombre d'heures hebdomadaires où l'enfant est en garderie (GCRQ1GA) ne s'applique que pour les enfants qui vont à la garderie (GCRQ1GA = 1). Dans les autres cas, on attribue le code 996 à cette question.

Là où la PMR n'a pas de conjoint de droit ou de fait qui vit dans le ménage, toutes les variables du conjoint/partenaire (sections sur la population active et l'éducation du conjoint/partenaire) sont codées « Sans objet ».

Non déclaré

Dans le cas du traitement qui se fait au bureau central, on code parfois comme « Non déclaré » la réponse à une question. On indique ainsi que la réponse à la question est inconnue. On attribue de tels codes pour quelques raisons:

- Dans le cadre de l'interview assistée par ordinateur, l'intervieweur pouvait entrer un code « Refus » ou « Ne sait pas », ainsi que nous l'avons expliqué plus haut. Le système IAO était souvent programmé en pareil cas pour sauter cette section particulière du questionnaire. En cas de refus, on supposait que les questions posées étaient délicates et qu'il était probable que le répondant ne veuille pas répondre à d'autres questions à ce sujet. Dans les cas « Ne sait pas », on partait du principe que le répondant n'était pas suffisamment informé pour répondre à d'autres questions. Dans le traitement des données, on a décidé que toutes les questions subséquentes se verraient attribuer un code « Non déclaré ». Nous entendons par-là que le répondant n'a pas répondu à la question. Nous ignorons même parfois si la question s'appliquait à ce dernier.
- Dans certains cas, il était impossible de remplir un questionnaire ni même parfois de le commencer. Par exemple, il a pu se produire une interruption ou le répondant a dit ne pas vouloir continuer. Si nous avons obtenu suffisamment de renseignements pour que le ménage soit considéré comme répondant, nous attribuons le code « Non déclaré » à toutes les questions restantes (et aux composantes dont nous n'avons pas encore commencé). Comme seule exception, si nous savions qu'une section ou qu'un questionnaire ne s'appliquait pas, nous avons codé ces éléments comme « Sans objet ».
- Dans une autre situation, nous avons recouru à des codes « Non déclaré » à la suite de vérifications de cohérence. Si nous décelions une erreur de cohérence entre des groupes de variables, une ou plusieurs des variables en question se voyaient attribuer le code « Non déclaré ».

Dans le cas des variables dérivées, si une ou plusieurs des variables d'entrée (aux variables dérivées) avaient reçu le code « Refus », « Ne sait pas » ou « Non déclaré », alors la variable dérivée était attribuée le code « Non déclaré ».

On a attribué un code « 9 » aux cas « Non déclaré ». Dans le cas d'une variable à un chiffre, le code est « 9 », pour une variable à deux chiffres, le code est « 99 » et pour une variable à trois chiffres, le code est « 999 », etc.

7.12 Nomenclature

Afin de faciliter l'interprétation des données par les utilisateurs, l'attribution des noms de variables et des valeurs est régie par certaines règles dans le système de documentation du fichier de microdonnées de l'Enquête sur les communautés de 2005.

Nous avons conçu une nomenclature pour chaque variable du fichier de données pour renseigner plus précisément l'utilisateur sur ces dernières. Tous les noms de variables sont d'au plus huit caractères, ce qui permet d'utiliser facilement ces désignations avec des progiciels d'analyse comme SAS ou SPSS. Les indicateurs « Persruk » et « Fieldruk » sont les exceptions à cette règle.

Format pour les noms de variables

Le **premier** caractère du nom de la variable désigne l'année de référence de l'Enquête sur les communautés :

- « C » indique l'année de référence de 1999,
- « D » indique l'année de référence de 2000,
- « E » indique l'année de référence de 2001,
- « F » indique l'année de référence de 2004, et
- « G » indique l'année de référence de 2005.

Les **deuxième et troisième** caractères désignent la section du questionnaire où la question a été posée ou dont la variable a été tirée. Veuillez consulter la section 7.13 pour les acronymes employés pour chaque section des questionnaires.

Le **quatrième** caractère du nom de la variable désigne l'unité de collecte ou l'unité à laquelle se rapporte la variable. Il y a cinq possibilités⁵ :

- « C » si la variable indique l'enfant,
- « P » si la variable indique la PMR,
- « S » si la variable indique le conjoint/partenaire,
- « H » si la variable indique le ménage, et
- « W » si la variable indique la variable pondérée.

Les **cinquième, sixième, septième et huitième** caractères du nom de la variable (par exemple GSDCQ2B et GLFPbD38) peuvent désigner :

- Q** désigne la variable qui est liée à une question directement posée dans le questionnaire.
- S** désigne un résultat calculé pour une des échelles utilisées dans le questionnaire.
- D** signifie que la variable a été calculée à partir d'au moins deux questions qui ont été posées dans le questionnaire ou qui ont été codées.
- I** signifie que la variable est un indicateur d'imputation.
- nnx** désigne l'identification de la question ou de la variable. Habituellement, « nn » est un numéro de séquence assigné à la variable; et « x » est un indicateur alphabétique de séquence pour une série de variables de type similaire.

7.13 Siglaison des sections du questionnaire

Le tableau qui suit indique les acronymes employés pour les diverses sections des questionnaires. Ces acronymes font partie du nom des variables du fichier de données. Ils occupent les deuxième et troisième positions du nom de la variable.

⁵ Il convient de noter que, bien qu'il y ait des variables pour une diversité d'unités d'analyse (PMR, conjoint/partenaire et ménage), on ne pourra produire que des « estimations pour les enfants » à partir du fichier. On peut se reporter aux caractéristiques de la PMR, du conjoint/partenaire et du ménage pour décrire les caractéristiques de l'enfant. Ainsi, il sera possible d'estimer le nombre d'enfants vivant dans un ménage à faible revenu ou dont la PMR a reçu une note élevée à l'échelle de la dépression. Il sera toutefois impossible de produire des estimations du nombre de ménages à faible revenu ou de PMR déprimées.

Acronyme	Section	Recueillie ou générée à partir de :
GE	Renseignements géographiques	Données d'échantillon.
HH	Ménage	Caractéristiques du logement.
MM	Variables recueillies au titre de la liste du ménage	Variables démographiques de base pour chaque membre du ménage; celles-ci figurent au fichier de données pour l'enfant, la PMR et le conjoint/partenaire.
DM	Variables démographiques calculées qui décrivent les conditions de vie de l'enfant	Données de la liste du ménage et de la grille des liens.
SD	Socio-démographiques	L'enfant au questionnaire de l'enfant et de la PMR et conjoint/partenaire au questionnaire de l'adulte.
HL	Santé des adultes	La PMR et conjoint/partenaire au questionnaire de l'adulte et de l'enfant au questionnaire de l'enfant.
CH	États chroniques des adultes	La PMR et conjoint/partenaire à la section sur la santé du questionnaire de l'adulte.
RS	Limitations d'activités	La PMR et conjoint/partenaire à la section sur la santé du questionnaire de l'adulte.
DP	Échelle de dépression	Questionnaire des parents (test d'échelle que subit la PMR).
ED	Éducation	L'enfant au questionnaire de l'enfant et la PMR et conjoint/partenaire au questionnaire de l'adulte.
LF	Population active	La PMR et conjoint/partenaire au questionnaire de l'adulte.
IN	Revenu	Revenu du ménage et revenu personnel de la PMR recueillis au questionnaire de l'adulte.
FN	Fonctionnement de la famille	Questionnaire de l'adulte (section administrée à la PMR ou au conjoint/partenaire).
AC	Activités	L'enfant au questionnaire de l'enfant.
BE	Comportement	L'enfant au questionnaire de l'enfant.
PR	Style parental	L'enfant au questionnaire de l'enfant.
CR	Garde des enfants	L'enfant au questionnaire de l'enfant.
PP	Test EVIP	L'enfant au questionnaire de l'enfant.
SF	Sécurité du voisinage	Questionnaire de l'adulte (section administrée à la PMR au conjoint/partenaire).
SP	Soutien social	Questionnaire de l'adulte (section administrée à la PMR au conjoint/partenaire).
SL	Sommeil	L'enfant au questionnaire de l'enfant.
WM	Qui suis-je?	Mesure directe.
KN	Connaissance des nombres	Mesure directe.
CC	Composante « Communautés »	L'enfant au questionnaire de l'enfant
WT	Poids	Poids faisant partie du plan de l'échantillon.

7.14 Exemples de noms de variables

Voici des exemples qui illustrent la nomenclature des variables au fichier de données :

Variable GLFSQ02	Désigne :
G	Indique l'année de référence de 2005
LF	Section sur la population active
S	Conjoint/partenaire
Q	Question directement posée dans le questionnaire
02	Deuxième question tirée de la section sur la population active du questionnaire de l'adulte

Variable GPRCS03	Désigne :
G	Indique l'année de référence de 2005
PR	Section sur le rôle parental
C	Enfant
S	Score – une interaction positive à l'échelle du rôle parental pour un enfant
03	Le numéro d'identification de la question

8.0 Contenu de l'enquête

L'Enquête sur les communautés a été conçue en fonction d'une approche écologique ou holistique pour mesurer le développement de l'enfant. L'enquête permet de saisir toute la diversité et la dynamique des facteurs ayant une incidence sur les enfants. Pour assurer une couverture adéquate de tous les domaines pertinents qui influent sur le développement de l'enfant, une consultation multidisciplinaire a eu lieu dès la conception de l'enquête. Le choix des priorités, des questions d'enquête et des domaines particuliers a résulté en grande partie d'un effort concerté soutenu par la contribution et les conseils :

- du groupe consultatif expert de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ), composé de spécialistes du développement de l'enfant et des sciences sociales;
- de ministères fédéraux;
- de représentants provinciaux et territoriaux responsables des programmes liés au développement de l'enfant.

Le groupe a recommandé que l'Enquête sur les communautés porte sur une large gamme de caractéristiques et de facteurs influant sur la croissance et le développement de l'enfant. On a recueilli une foule de renseignements au sujet de l'enfant, de son ou ses parents, des caractéristiques de la famille et du quartier ainsi que de l'école de l'enfant et sa vie à l'école. La présente section donne un aperçu du contenu de chacune des sections du questionnaire ayant servi à la collecte des données de l'Enquête sur les communautés ainsi qu'une description des différentes échelles utilisées dans le cadre de l'Enquête sur les communautés.

8.1 Composantes de l'enquête

L'Enquête sur les communautés est subdivisée en plusieurs composantes, lesquelles sont décrites au chapitre 6.0, Collecte des données. Un sommaire de chaque composante est présenté ci-dessous.

Ménage	Cette composante constitue la première partie de l'interview. On recueille, pour la liste des membres du ménage, des renseignements démographiques de base sur chacun des membres et des données sur les liens qui l'unissent à tous les autres membres du ménage.
Adulte	Questions relatives à la personne la mieux renseignée (PMR) et son conjoint. La composante des adultes est remplie une seule fois, même lorsque le ménage comprend deux enfants.
Enfant	Les questions relatives à l'enfant sélectionné sont posées à la PMR.
Évaluations directes	On effectue plusieurs évaluations directes relativement aux enfants; celles-ci sont décrites au chapitre 6.0.

8.2 Variables démographiques

Les variables démographiques sont recueillies dans le questionnaire du ménage. À partir de ce questionnaire, certains renseignements démographiques de base (p. ex., l'âge, le sexe, l'état matrimonial) sont recueillis pour tous les membres du ménage de l'enfant. Ce questionnaire vise en outre à recueillir des données sur la grille des liens, c'est-à-dire les liens entre tous les membres du ménage et la PMR. Ces données ont permis de créer un vaste ensemble de

variables pour décrire la situation familiale de l'enfant. La plupart de ces variables dérivées sont essentielles à l'analyse des données de l'Enquête sur les communautés et sont décrites au chapitre 4.0.

Il a fallu procéder à toute une série de vérifications des données recueillies. Voici des exemples des vérifications effectuées :

- un parent biologique doit avoir au moins 12 ans de plus (et au maximum 55 ans de plus) que l'enfant biologique;
- la différence d'âge entre mari et femme doit être inférieure à 29 ans.

8.3 Questionnaire de l'adulte

Éducation (parent)

La section sur le niveau de scolarité est remplie pour la PMR et son conjoint/partenaire. L'objectif est de recueillir des renseignements sur le plus haut niveau de scolarité atteint.

Certaines recherches ont montré qu'il existe un lien entre le niveau de scolarité de la mère, le milieu familial et le développement de l'enfant.

Activité sur la population active

La stabilité d'emploi a un effet sur le milieu familial, tant sur le plan du revenu que sur celui du niveau de stress. Certaines recherches montrent que le chômage des parents peut avoir un effet néfaste sur les résultats des enfants.

La section relative à l'activité sur la population active est remplie à la fois pour la PMR et son conjoint/partenaire. L'objectif principal de cette section est de déterminer la stabilité d'emploi, puisque celle-ci constitue un indicateur de la constance du revenu d'emploi. Les questions portent sur les périodes d'absence du travail, les raisons de l'absence la plus récente, le nombre d'heures de travail et l'organisation du travail (par exemple, travail par quarts) au cours de l'année précédente. On a recueilli des renseignements portant sur l'emploi actuel ou l'emploi le plus récent de la PMR et de son conjoint/partenaire.

Une description complète de l'emploi actuel ou de l'emploi le plus récent est consignée. La branche d'activité et la profession sont codées selon le Système de classification des industries de l'Amérique du Nord (SCIAN) de 1997 et la Classification type des professions (CTP) de 1991.

Variables dérivées relatives à l'activité sur la population active

Plusieurs variables dérivées relatives à l'activité sur le marché du travail ont été créées pour la PMR et son conjoint/partenaire. Parmi ces variables figurent les suivantes :

- GLFPcD5A / GLFSCD5A : code du SCIAN pour l'emploi actuel de la PMR / de son conjoint
- GLFPCD6A / GLFScD6A : code de la CTP 1991 pour l'emploi principal de la PMR / de son conjoint
- GLFPcD7A / GLFSCD7A : code de la classification type des industries pour l'emploi actuel - regroupé
- GLFPcD8A / GLFSCD8A : code de la classification type des professions pour l'emploi actuel - regroupé

Revenu

Dans la section de l'enquête portant sur le revenu, on recueille des renseignements sur les sources et le montant du revenu pour chacun des ménages. Il y a également quelques questions portant sur la perception qu'ont la PMR et son conjoint de leur situation financière. Ces renseignements constituent un indicateur de la situation économique de la famille, une composante essentielle du cadre de vie de l'enfant.

Puisque le revenu familial est une variable importante de bon nombre d'études sur le développement de l'enfant, on a imputé une valeur au revenu du ménage en l'absence de réponses à ces questions. Le chapitre 4.0 présente une explication détaillée de la procédure d'imputation du revenu.

On a créé deux variables dérivées (GINHD04A et GINHD05A) afin de comparer le revenu du ménage au seuil de faible revenu (SFR). Les seuils de faible revenu servent à distinguer les unités familiales qui sont à « faible revenu » des « autres » groupes. Une unité familiale est considérée à faible revenu lorsque son revenu est inférieur à la valeur du seuil correspondant à sa taille d'unité familiale et à la taille de sa communauté. Toute famille ayant un revenu égal ou supérieur au seuil de faible revenu est considérée dans la catégorie « autre »⁶. La variable GINHD03A donne la valeur du SFR par zone géographique.

Santé (adulte)

Cette section comporte des questions adressées aux PMR et à leur conjoint au sujet de leur état général de santé, les états chroniques et les limitations.

États chroniques

On demande aux PMR et à leur conjoint s'ils souffrent d'états chroniques (p. ex., allergies, asthme, hypertension). Une variable dérivée (GCHPD01 ou GCHSD01) indique une réponse affirmative, c'est-à-dire que les répondants ont déclaré au moins un état chronique.

Limitations d'activités

On pose aux PMR et à leur conjoint une série de questions sur les limitations de leurs activités à la maison, au travail, à l'école, etc. Une variable dérivée (GRSPD01 ou GRSSD01) permet de déterminer si la PMR ou son conjoint ont déclaré des limitations d'activités.

Échelle de dépression

Les PMR répondent aux questions se rapportant à l'échelle de dépression (GDPPS01) dans le cadre du questionnaire de l'adulte.

Fonctionnement de la famille

Le but de cette section est de fournir une évaluation globale du fonctionnement de la famille et une indication de la qualité des relations familiales. Ces questions sont posées à la PMR ou à son conjoint.

Sécurité du quartier

Cette section permet de recueillir des renseignements sur la satisfaction des répondants quant à leur quartier en tant que milieu de vie pour les enfants, y compris leurs perceptions relativement à l'importance des dangers et des problèmes, à la cohésion sociale ou sentiment de « voisinage ». Deux échelles ont été créées dans cette section : celle de la sécurité du quartier (GSFHS5), qui mesure le degré de sécurité du quartier tel que le perçoivent les répondants, et l'échelle du voisinage (GSFHS6), qui mesure le degré de cohésion sociale dans le quartier.

Soutien social

L'objet de cette section est de recueillir des renseignements sur le niveau de soutien que les PMR pensent avoir de la part de leurs amis, des membres de leur famille et de leur collectivité. Ces questions sont posées à la PMR ou à son conjoint.

Caractéristiques sociodémographiques

La section sur les caractéristiques sociodémographiques vise à recueillir des renseignements sur l'immigration, l'origine ethnique et le profil linguistique des membres du ménage. Ces données

⁶ Pour de plus amples renseignements sur les mesures utilisées par Statistique Canada en matière de faible revenu, veuillez consulter *Seuils de faible revenu de 1994 à 2003* et *Mesures du faible revenu de 1992 à 2001*, N° 75F0002MIF, 2004002 au catalogue.

permettent l'analyse des diverses composantes de la population canadienne et l'identification des personnes appartenant à une minorité visible. La section comporte en outre une question sur l'appartenance religieuse. Ces questions se rapportent à la PMR, à son conjoint et à l'enfant.

8.4 Questionnaire de l'enfant

Mesures directes

L'objet de cette section est de préparer le terrain pour les tests de mesures directes qu'on fera passer aux enfants. Si les enfants ne peuvent pas y répondre (par exemple, s'il ne parlent ni l'anglais ni le français ou s'ils sont daltoniens), on ne leur fera pas passer les tests.

Santé (enfant)

Cette section vise à recueillir des renseignements sur l'état de santé physique de l'enfant – état général de santé, blessures, limitations et états chroniques – de même que sur le recours aux services de santé.

Alphabétisation

Cette section permet de mesurer le contact qu'ont les enfants avec les livres, leur intérêt pour la lecture et les activités d'apprentissage auxquelles les parents s'adonnent avec leurs enfants. La section porte principalement sur la stimulation des jeunes enfants reçue à la maison.

On a ajouté une question sur les activités de calcul, conformément aux recommandations de collègues de D^r Robbie Case, l'auteur du test sur la connaissance des nombres, selon qui les capacités de calcul constituent un facteur déterminant en ce qui a trait à l'apprentissage et à l'alphabétisation. On a ajouté cette question pour obtenir des renseignements sur les activités de calcul des enfants susceptibles de confirmer les résultats du test de connaissance des nombres.

Communications

Ces questions sont adaptées à partir de la Competent Children Study de la Nouvelle-Zélande. Elles portent sur les capacités de l'enfant de comprendre les messages oraux, de transmettre un message à autrui et de communiquer verbalement.

Activités

Cette section permet de mesurer la participation des enfants à diverses activités non scolaires ainsi qu'une indication quant à l'emploi du temps de l'enfant, ses intérêts personnels et son degré d'interaction avec ses pairs.

Communauté

Cette section vise à déterminer l'existence des ressources éducatives et des ressources récréatives dans la communauté comme, par exemple, des musées, des centres communautaires et les raisons particulières pour lesquelles l'enfant ne participe pas, par exemple, l'inaccessibilité ou le coût.

Comportement

L'objectif de cette section est d'évaluer certains aspects du comportement de l'enfant et à mesurer la fréquence de comportements tels que l'hyperactivité et l'agressivité physique. Les échelles établies à partir de ces questions sont décrites de façon détaillée plus tard dans ce chapitre.

Comportement positif

Cette section vise à évaluer le comportement positif de l'enfant, notamment la persévérance et l'autonomie. Selon la Competent Children Study de la Nouvelle-Zélande, la persévérance et l'autonomie figurent parmi l'ensemble des compétences constituant de bons indicateurs du rendement général de l'enfant.

Les questions sont inspirées de l'étude néo-zélandaise et des questions de l'ELNEJ se rapportant au comportement des enfants d'autres groupes d'âge.

Sommeil

Ces questions portent sur le nombre d'heures de sommeil quotidien que l'enfant dort en moyenne, incluant les siestes de jour et celles de nuits.

Rôle parental

On estime que le rôle parental exerce une influence importante sur le comportement et le développement de l'enfant. L'objectif de cette section est de mesurer certains comportements parentaux. On a créé des échelles à partir des questions de cette section.

On a demandé certaines questions aux PMR ayant un conjoint/partenaire à la maison pour déterminer jusqu'à quel point ils s'entendent au sujet des décisions parentales. Cette question s'apparente à des questions liées à l'échelle de Strayhorn et Weidham dont sont inspirées les autres questions sur le rôle parental.

Antécédents familiaux

Cette section vise à fournir des renseignements sur la situation familiale de l'enfant : parents mariés ou non, séparés ou divorcés, âge de l'enfant au moment de la séparation ou du divorce, etc.

Garde d'enfants

Cette section fournit des renseignements de base sur les services de garde actuels des enfants pendant que les parents travaillent ou étudient, ainsi que des données sur les services de garde antérieurs. Les données permettent notamment de mesurer le temps passé par l'enfant en garderie et le genre de service de garde utilisé pour chaque enfant. On recueille également des renseignements sur le nombre de changements de services de garde qu'a vécus l'enfant ainsi que sur les raisons de ces changements au cours des 12 derniers mois. Ces questions permettent aussi de déterminer s'il s'agit d'une garderie à but lucratif ou sans but lucratif et s'il s'agit de services de garde en milieu familial agréés ou non agréés ainsi que le ratio « travailleurs de services de garde : enfants ».

Caractéristiques sociodémographiques

Ces questions visent à recueillir des renseignements sociodémographiques sur les enfants sélectionnés, par exemple leur origine ethnique et leur pays d'origine.

8.5 Échelles de l'enquête

Pour certains des concepts dont la mesure a été jugée importante pour l'Enquête sur les communautés, on a décidé qu'il conviendrait mieux d'utiliser une échelle. Une échelle est constituée simplement d'un groupe de questions dont les réponses mises ensemble permettent de mesurer certains concepts.

Par exemple, à partir du questionnaire de l'enfant, on a déterminé qu'il serait important d'évaluer certains comportements parentaux. L'échelle vise à mesurer quatre concepts ou facteurs liés au rôle parental : l'interaction positive, l'inefficacité parentale la cohérence des pratiques parentales et la rationalité des pratiques parentales.

Pour chaque facteur mesuré au moyen d'une échelle, on calcule un score (note). Le score attribué à un facteur particulier peut servir à classer les personnes. Par exemple, pour les enfants ayant les scores les plus élevés sur l'échelle d'interaction positive, les personnes la mieux renseignées (PMR) ont déclaré avoir plus de rapports positifs avec l'enfant (ils riaient ensemble davantage, l'enfant était félicité plus souvent, etc.). Le score pour un facteur particulier est généralement fondé sur une série de questions, étant donné qu'une seule question ne peut mesurer le concepts ou le facteurs avec suffisamment de précision.

Au moment d'élaborer l'Enquête sur les communautés, il a été décidé d'emprunter les mêmes échelles qui sont utilisées dans ELNEJ pour les enfants âgés de cinq ans. Pour de plus amples renseignements visant la méthode de calculer les scores, veuillez faire référence au chapitre 9.0 du Guide de l'utilisateur des microdonnées de l'ELNEJ, cycle 5. Le reste du chapitre présente une description détaillée des sources des échelles de l'Enquête sur les communautés.

8.5.1 Échelle de comportement

L'objectif de cette échelle est d'évaluer les comportements suivants chez les enfants:

- Les problèmes de conduite
- Hyperactivité
- Troubles émotifs
- L'agression indirecte

Concepts théoriques

Les concepts théoriques utilisés pour l'analyse des facteurs sont présentés ci-dessous. Les échelles réelles émanant de l'analyse diffèrent de ces concepts.

Troubles de conduite

Questions GBECQ6G, GBECQ6X, GBECQ6AA, GBECQ6FF, GBECQ6JJ et GBECQ6NN tirées de l'Étude sur la santé des jeunes ontariens (ESJO).

Hyperactivité

Questions GBECQ6B, GBECQ6I, GBECQ6P, GBECQ6S et GBECQ6W tirées de l'ESJO, et questions GBECQ6QQ et GBECE69B tirées de l'Enquête longitudinale de Montréal.

Troubles émotifs

Questions GBECQ6F, GBECQ6K, GBECQ6Q, GBECQ6V, GBECQ6CC, GBECQ6MM et GBECQ6RR tirées de l'ESJO. L'anxiété est mesurée par les questions de l'ELNEJ tirées des questions de l'ESJO se rapportant aux troubles émotifs (GBECQ6F, GBECQ6Q, GBECQ6V et GBECQ6CC).

Agressivité indirecte

Questions GBECQ6J, GBECQ6R, GBECQ6Z, GBECQ6LL et GBECQ6TT tirées des travaux des chercheurs finlandais Lagerspetz, Bjornqvist et Peltonen.

Les items qui ont servi à mesurer ces différents concepts de comportement sont présentés dans le tableau ci-dessous.

FACTEUR	SCORE	ITEMS
Hyperactivité – inattention	GBECdS6A	GBECQ6B, GBECQ6I, GBECQ6P, GBECQ6S, GBECQ6W, GBECQ6QQ, GBECE69B
Troubles émotifs – anxiété	GBECdS8A	GBECQ6F, GBECQ6K, GBECQ6Q, GBECQ6V, GBECQ6CC, GBECQ6MM, GBECQ6RR
Agressivité physique – trouble de conduite	GBECS09	GBECQ6G, GBECQ6X, GBECQ6AA, GBECQ6FF, GBECQ6JJ, GBECQ6NN
Agressivité indirecte	GBECS10	GBECQ6J, GBECQ6R, GBECQ6Z, GBECQ6LL, GBECQ6TT

8.5.2 Échelle de dépression

Objectifs et aperçu

L'échelle de dépression a été administrée à la PMR dans le questionnaire du parent. Les questions rattachées à cette échelle (GDPPQ12A à GDPPQ12L) constituent une version réduite de l'échelle de dépression (CES-D) de 20 questions mise au point par L.S. Radloff du Centre d'étude épidémiologique du *National Institute of Mental Health* des États-Unis. Cette échelle sert à mesurer la fréquence des symptômes dans le grand public. On mesure la fréquence et la gravité des symptômes associés à la dépression pendant la semaine précédente. L'échelle a été réduite à 12 questions par M. Boyle, de l'Hôpital Chedoke-McMaster de l'Université McMaster.

L'objectif de cette échelle est de recueillir des renseignements sur la santé mentale des répondants, surtout en ce qui a trait aux symptômes de dépression. Plusieurs membres du groupe consultatif expert de l'ELNEJ ont laissé entendre que la meilleure façon de procéder est de mesurer un aspect particulier de la santé mentale de la PMR, au lieu de tenter de mesurer son état général de santé mentale. On a proposé que cette section soit centrée sur la dépression pour les raisons suivantes : il s'agit d'un état prévalent; il a été démontré que la dépression d'un parent a des répercussions sur les enfants; les recherches actuelles sur ce sujet sont généralement fondées sur des groupes de démonstration et non sur des échantillons de population; et on croit que l'adoption de politiques dans ce domaine peut améliorer la situation.

Questions rattachées à l'échelle de dépression

L'échelle de dépression comporte douze questions et chacune d'elles est associée à quatre catégories de réponse. Afin que la plus faible valeur du score soit 0, la valeur des catégories de réponse a été réduite de 1 lors du calcul du score. On a également inversé les catégories de réponse pour les questions ayant un poids factoriel négatif (GDPPQ12F, GDPPQ12H, GDPPQ12J). Le score total (GDPPS01) peut donc varier de 0 à 48, un score élevé dénotant la présence de symptômes de dépression.

8.5.3 Échelle de fonctionnement de la famille

Objectifs et aperçu

Les questions sur le fonctionnement de la famille (GFNHQ01A à GFNHQ01L) ont été élaborées par les chercheurs de l'Hôpital Chedoke-McMaster de l'Université McMaster et sont largement utilisées tant au Canada qu'à l'étranger. Il s'agit d'une échelle qui sert à mesurer différents aspects du fonctionnement de la famille (la résolution de problèmes, la communication, les rôles, l'engagement affectif, la sensibilité affective et la maîtrise du comportement).

La question GFNHQ01M, tirée du suivi de l'Étude sur la santé des jeunes ontariens, a été ajoutée à l'échelle originale pour déterminer si la consommation d'alcool avait des répercussions sur la dynamique globale de la famille. Elle n'a cependant pas été retenue dans l'analyse de l'échelle.

L'objectif de cette échelle est d'obtenir une évaluation globale du fonctionnement de la famille et une indication de la qualité des relations entre les parents ou partenaires. Pour cette raison et étant donné le nombre restreint de questions, on n'a pas tenté de mesurer les différents aspects du fonctionnement de la famille.

D'autres enquêtes ont montré que les relations entre les membres de la famille ont des répercussions considérables sur les enfants. Par exemple, les résultats de l'Étude sur la santé des jeunes ontariens indiquent un lien important entre le dysfonctionnement familial et certains problèmes de santé mentale chez l'enfant.

Administration de l'échelle de fonctionnement de la famille

L'échelle de fonctionnement de la famille a été administrée à la PMR ou à son conjoint/partenaire dans le cadre du questionnaire du parent. L'échelle comporte 12 questions et chacune d'elles est associée à quatre catégories de réponse. Pour que la plus faible valeur du score soit 0, on a réduit la valeur des catégories de réponse de 1 lors du calcul du score. On a également inversé les catégories de réponse pour les questions ayant un poids factoriel négatif (GFNHQ01A, GFNHQ01C, GFNHQ01E, GFNHQ01G, GFNHQ01I et GFNHQ01K). Le score total (GFNHS01) peut donc varier de 0 à 36, un score élevé dénotant un dysfonctionnement de la famille.

8.5.4 Échelle de sécurité du quartier

Objectifs et aperçu

Cette échelle vise à réunir des renseignements sur la satisfaction du répondant à l'égard de son quartier en tant de milieu de vie pour les enfants, y compris sur les perceptions du niveau de risque et de problèmes ainsi que de cohésion sociale ou « sentiment de voisinage ». Selon des recherches récentes menées par Jacqueline Barnes du Judge Baker Children's Centre de l'Université Harvard à Boston, la peur du danger ressentie par les parents et les perceptions de désordre social au sein du quartier ont une incidence sur leur attachement au quartier et sur leurs stratégies en matière de discipline.

GSFHQ01 à GSFHQ06E : Ces questions couvrent la durée de résidence dans le quartier, la satisfaction à l'égard du quartier en tant que milieu de vie pour les enfants, la sécurité, la cohésion sociale et les problèmes du quartier. Il s'agit d'une version révisée des sections particulières du Simcha-Fagan Neighbourhood Questionnaire dont s'est servie Jacqueline Barnes dans ses travaux sur des quartiers de Boston et de Chicago. Ce questionnaire a été modifié, en consultation avec Jacqueline Barnes, à la lumière de l'analyse des facteurs des sections.

Les items rattachés à chaque facteur sont décrits au tableau suivant.

FACTEUR	SCORE	ITEMS
Sécurité du quartier	GSFHS5	GSFHdQ5A, GSFHdQ5B, GSFHdQ5C
Voisinage	GSFHS6	GSFHQ06A, GSFHQ06B, GSFHQ06C, GSFHQ06D, GSFHQ06E

Veuillez noter que dans l'Enquête sur les communautés de 2000 et celui de 2001, il n'y avait qu'une échelle, l'échelle de voisinage.

8.5.5 Échelles du rôle parental

Objectifs et aperçu

L'objectif de cette échelle est de mesurer certaines pratiques parentales. Plus précisément, deux échelles ont été utilisées. La première a été conçue afin de mesurer l'interaction positive, l'hostilité ou l'inefficacité et la cohérence des parents envers les enfants. La deuxième échelle vise à mesurer les pratiques parentales provoquant ou non l'aversion.

Les questions GPRCQ01 à GPRCQ18 et GPRCQ21 à GPRCQ24 sur l'interaction positive, sur l'hostilité ou l'inefficacité et sur la cohérence ont été fournies par M. Boyle de l'Hôpital Chedoke-McMaster, à partir des travaux de Ken Dodge (Université Vanderbilt) et d'une adaptation de la Parent Practices Scale de Strayhorn et Weidman.

Calcul des scores des pratiques parentales

Une fois que l'on a analysé les structures factorielles et établi la liste des items associés à chaque facteur, on a calculé les scores des différents facteurs. Pour le calcul des scores, on a retranché 1 de la valeur de chacun des items afin que la plus faible valeur possible pour le score soit 0. Pour chacun des quatre facteurs observés, un score de 0 indique :

- l'absence d'interactions positives pour le facteur d'interaction positive;
- l'absence d'interactions hostiles ou inefficaces pour le facteur d'hostilité/inefficacité;
- l'absence de pratiques parentales cohérentes pour le facteur de cohérence;
- la présence d'interactions punitives ou de pratiques provoquant l'aversion pour le facteur d'hostilité ou d'inefficacité.

L'analyse factorielle a permis de définir quatre facteurs pour ce groupe d'âge : interaction positive (GPRCS03), hostilité (GPRCS04), cohérence (GPRCS05) et pratiques punitives/provoquant l'aversion (rationalité) (GPRCS06). Les items rattachés à chaque facteur sont présentés ci-dessous.

FACTEUR	SCORE	ITEMS
Interaction positive	GPRCS03	GPRCQ01, GPRCQ02, GPRCQ03, GPRCQ06, GPRCQ07
Inefficacité	GPRCS04	GPRCQ04, GPRCQ08*, GPRCQ09, GPRCQ13, GPRCQ14, GPRCQ15, GPRCQ18
Cohérence	GPRCS05	GPRCQ10, GPRCQ11, GPRCQ12*, GPRCQ16*, GPRCQ17*
Rationalité	GPRCS06	GPRCQ21, GPRCQ22*, GPRCQ23, GPRCQ24*

* indique que la valeur de l'item a été inversée dans le calcul du score. Dans l'Enquête sur les communautés de 2000 et celui de 2001, GPRCQ21 et GPRCQ23 ont été renversées plutôt que GPRCQ22 et GPRCQ24.

8.5.6 Échelle du soutien social

Objectifs et aperçu

L'échelle initiale comportait 24 items tirés du Social Provisions Model de Robert Weiss, lequel décrit six fonctions ou « dispositions » sociales distinctes qu'il est possible d'acquérir grâce aux relations avec autrui. Compte tenu de la taille de l'échelle et suivant les recommandations de M. Boyle de l'Hôpital Chedoke-McMaster, on a utilisé dans le cadre de l'enquête une version inspirée du projet Partir d'un bon pas pour un avenir meilleur du gouvernement de l'Ontario. Cette échelle permet de mesurer l'encadrement, les alliances fiables et l'attachement.

Les valeurs associées aux items GSPHQ01A, GSPHQ01E, GSPHQ01F et GSPHQ01I ont été inversées dans le calcul du score. Toutes les valeurs ont été recodées, passant de 1-4 à 0-3. Le score final varie de 0 à 24, un score élevé indiquant un degré élevé de soutien social. La structure factorielle est identique à celle employée dans le cycle 1 de l'ELNEJ.

Les items rattachés à chaque facteur sont décrits au tableau suivant.

FACTEUR	SCORE	ITEMS
Soutien social	GSPHS01	GSPHQ01A, GSPHQ01B, GSPHQ01C, GSPHQ01D, GSPHQ01E, GSPHQ01F, GSPHQ01H, GSPHQ01I.

9.0 Qualité des données

9.1 Taux de réponse

Le tableau ci-dessous résume les taux de réponse de l'Enquête sur les communautés de 2005.

Communautés	Taille de l'échantillon	Taille de l'échantillon exclus hors du champ de l'enquête	Répondants	Non-répondants	Incapable de contacter	Taux de réponse (%)
South Eastman (Manitoba)	582	567	475	77	15	83,8
Hampton (Nouveau-Brunswick)	697	691	574	73	44	83,1
Abbotsford (Colombie-Britannique)	582	565	462	65	38	81,8
Mississauga (Ontario)	668	640	539	51	50	84,2
Niagara Falls (Ontario)	721	694	582	66	46	83,9
Saskatoon (Saskatchewan)	584	571	467	48	56	81,8
Montréal (Québec)	575	572	466	50	56	81,5
Total	4 409	4 300	3 565	430	305	82,9

9.2 Erreurs relatives à l'enquête

Des sept communautés à l'étude, quatre (South Eastman, Abbotsford, Saskatoon et Montréal) ont leurs estimations calculées à partir d'un échantillon d'enfants. Des estimations légèrement différentes auraient pu être obtenues si un recensement complet avait été effectué en reprenant le même questionnaire et en faisant appel aux mêmes intervieweurs, surveillants, méthodes de traitement, etc. que ceux effectivement utilisés dans l'enquête. L'écart entre les estimations découlant de l'échantillon et celles que donnerait un dénombrement complet dans des conditions semblables est appelé erreur d'échantillonnage de l'estimation.

Pour les trois communautés restantes (Hampton, Mississauga et Niagara Falls) un recensement total des enfants à la maternelle avait été effectué. L'erreur d'échantillonnage est donc impossible pour ces cas.

Des erreurs qui ne sont pas liées à l'échantillonnage peuvent se produire à presque toutes les étapes des opérations d'enquête. Les intervieweurs peuvent avoir mal compris les instructions, les enquêtés peuvent se tromper en répondant aux questions, les réponses peuvent être mal saisies sur le questionnaire et des erreurs peuvent survenir lors du traitement et de la totalisation des données. Ces erreurs sont toutes des exemples d'erreurs non dues à l'échantillonnage.

9.3 Base de sondage

La base de sondage, qui consistait en une liste de tous les enfants inscrits à la maternelle, a été fournie par les commissions scolaires de sept communautés participant à notre étude. Dans tous les cas, les premières listes que nous avons reçues renfermaient plusieurs cas où des

renseignements étaient manquants tels que le numéro de téléphone ou l'adresse de l'enfant ou bien il y avait plusieurs dossiers pour un même enfant.

Pour les cas où des renseignements étaient manquants ou fournis en double, on communiquait avec les commissions scolaires qui avaient envoyé ces fichiers. Tous les cas en double ont été résolus et, lorsque c'était possible, les renseignements manquants ont été obtenus.

Étant donnée que les fichiers de la base de sondage étaient obtenus de commissions, districts ou divisions scolaires différentes provenant de sept communautés différentes, la qualité des fichiers variait considérablement d'une commission à l'autre. Cette observation avait été soulevée lorsque les cas renfermant des renseignements manquants ou en double avaient été identifiés. Certaines communautés fournissaient bien plus de cas problématiques devant être résolus que d'autres.

La base de sondage pouvait être considérée à jour au moment où elle avait été fournie, c'est-à-dire à l'automne 2004. Cependant, la base de sondage ne peut être considérée comme étant mise à jour durant la période de collecte. Il est possible que les enfants inscrits à la maternelle à l'automne, c'est-à-dire au moment où la base avait été obtenue, n'étaient pas inscrits à cette commission scolaire lors de la période de collecte. Ces enfants étaient donc hors du champ de l'enquête. Il y a aussi des enfants qui changent d'adresse sans toutefois changer d'école durant l'année scolaire. Le cas échéant, les renseignements sur la personne-ressource sur la base de sondage n'étaient pas à jour. Vu qu'il est impossible de repérer ces cas avant le début de la période de collecte, la qualité de la base se voit alors touchée.

Il a aussi été souligné que plusieurs enfants faisant partie de la base de sondage ne fréquentaient pas l'école dans la même communauté où ils vivaient. Étant donné que tous les enfants de la communauté qui étaient inscrits à la maternelle faisaient partie de la base de sondage peu importe leur lieu de résidence, les enfants qui n'habitaient pas l'une des sept communautés à l'étude étaient tout de même inclus dans cette base. Les enfants qui ont été inscrits à la maternelle après que la base de sondage avait été envoyée à Statistique Canada n'étaient pas admissibles à la sélection d'échantillons et, par conséquent, ne sont pas représentés dans l'estimation. Ceci comprend les familles qui ont déménagé dans ces communautés plus tard dans l'année. Ces situations auront certainement un impact lorsque viendra le temps de comparer les résultats et les estimations à d'autres enquêtes ayant des frontières géographiques plus fidèles des communautés étudiées.

9.4 Répondants hors du champ de l'enquête

Lors de la création de la base de sondage pour les enfants admissibles à la composante Communauté de l'enquête, on utilise plusieurs étapes afin d'assurer la plus haute qualité de renseignements possible. Dès l'automne 2004, les conseils scolaires et districts participants ont fourni à Statistique Canada les listes des enfants inscrits à la maternelle dans leurs écoles. Après plusieurs étapes pour éliminer les enregistrements incorrects ou en double, la base de sondage pour l'échantillonnage est finalisée. Cependant, la base utilisée pour sélectionner l'échantillon des communautés ne peut être considérée comme étant à jour qu'au moment où elle a été fournie, c'est-à-dire en automne 2004.

Ainsi, cette base peut ne plus être à jour au moment de la collecte des données, car il est en effet possible que des enfants inscrits en automne à la maternelle dans l'une des écoles choisies ne fréquentent plus cette école durant la période de collecte. Certains enfants peuvent avoir déménagé hors des communautés concernées; ils seraient donc hors du champ de l'enquête au moment de la collecte. D'autres enfants restés au sein des communautés concernées, mais ayant changé d'école ont également été retirés de l'échantillon parce que l'application n'est pas capable de prendre en compte de tels cas. Par ailleurs, les enfants qui ont emménagé dans les communautés concernées et se sont inscrits à l'école après l'automne n'étaient pas admissibles dans la sélection de l'échantillon et n'ont donc aucune chance d'être interviewés. Ces cas de

sous-dénombrement ne peuvent être mesurés dans le processus de cette enquête. Les inférences sur la population exclue ne peuvent être faites sur la base du nombre d'unités hors du champ de l'enquête que si l'hypothèse de la migration nette est proche de zéro.

Parmi les 4 409 enfants de l'échantillon, il y a eu 109 cas d'enfants considérés hors du champ de l'enquête ou retirés de l'échantillon. Ces enfants ont été exclus de la collecte des données puisqu'ils ne fréquentent plus l'école dans laquelle ils avaient été échantillonnés. En effet, dès qu'on a eu la confirmation que des enfants ne fréquentent plus l'école qui a donné leur nom permettant de les échantillonner ultérieurement, on les a codés comme étant des unités hors du champ de l'enquête. Toutefois, même si aucune interview n'a été menée auprès de ces enfants, on a recueilli certains renseignements au point de contact afin d'évaluer leur statut. Les enfants hors du champ de l'enquête ne font pas partie de ce fichier de données.

10.0 Lignes directrices pour la totalisation, l'analyse et la diffusion de données

Ce chapitre de la documentation renferme un aperçu des lignes directrices que doivent respecter les utilisateurs qui totalisent, analysent, publient ou diffusent des données calculées à partir des fichiers de microdonnées de l'enquête. Ces lignes directrices devraient permettre aux utilisateurs de microdonnées de produire les mêmes chiffres que ceux produits par Statistique Canada, tout en étant en mesure d'obtenir des chiffres actuels non-publiés de façon conforme à ces lignes directrices établies.

10.1 Lignes directrices pour la pondération de l'échantillon en vue de la totalisation

Le plan d'échantillonnage utilisé pour l'Enquête sur les communautés est autopondéré dans chacune des communautés. Chaque communauté a différents poids d'échantillonnage afin de refléter les différentes fractions de sondage. Certaines communautés n'ont pas été échantillonnées et un recensement des enfants inscrits à la maternelle dans ces communautés a été effectué. Lorsqu'ils produisent des estimations simples, y compris de tableaux statistiques ordinaires, les utilisateurs doivent appliquer le poids démographique approprié (WGT). Même dans les communautés où un recensement a été effectué, un poids peut exister pour corriger les erreurs de sous-dénombrement attribuables au fait que certaines familles n'ont pas répondu à l'enquête.

Si les poids appropriés ne sont pas utilisés, les estimations calculées à partir des fichiers de microdonnées ne peuvent être considérées comme représentatives de la population visée par l'enquête et ne correspondront pas à celles produites par Statistique Canada. En fait, le poids attribué à chaque enfant reflète le nombre d'enfants représentés par un répondant donné. Pour toute analyse comportant un volet de corrélation ou toute autre méthode statistique nécessitant un niveau de confiance, il est recommandé d'utiliser un poids d'échantillonnage. Ce poids est obtenu en multipliant le poids démographique (WGT) par la taille de l'échantillon, puis en divisant le total ainsi obtenu par la population totale estimée. Nous obtenons alors un poids moyen de 1 et la somme des poids est égale à la taille de l'échantillon. L'avantage d'utiliser le poids normalisé est qu'il permet d'éviter une surestimation de la confiance (laquelle est très sensible à la taille de l'échantillon), tout en maintenant les mêmes distributions que celles obtenues avec le poids démographique.

10.1.1 Définitions de types d'estimations : catégoriques et quantitatives

Il est à souligner que le fichier des Communautés a été établi de façon à ce que l'enfant constitue l'unité d'analyse. Le poids qui paraît sur chaque enregistrement (WGT) est un poids correspondant à un « enfant ». Il est donc impossible de calculer des estimations concernant les parents ou les familles à partir du fichier de données des Communautés. Vous trouverez plus de renseignements sur les unités d'analyse à la section 4.1 du présent document.

Avant de discuter de la façon dont on peut totaliser et analyser les données des Communautés, il est utile de décrire les deux principaux types d'estimations ponctuelles des caractéristiques de la population qui peuvent être produites à partir du fichier de microdonnées des Communautés.

Estimations catégoriques

Les estimations catégoriques sont des estimations du nombre ou du pourcentage de membres de la population visée par l'enquête possédant certaines caractéristiques ou

faisant partie d'une catégorie définie. Le nombre d'enfants nés avant terme ou la proportion d'enfants qui sont en excellente santé constituent des exemples de telles estimations. Une estimation du nombre de personnes possédant une certaine caractéristique peut aussi être désignée une estimation d'un total.

Exemples de questions catégoriques :

Q : En général, diriez-vous que la santé de (l'enfant) est :
R : Excellente / Très bonne / Bonne / Passable / Mauvaise

Q : Est-ce que (l'enfant) est né avant, après ou à terme?
R : Avant / Après / À terme

Estimations quantitatives

Les estimations quantitatives sont des estimations de totaux ou de moyennes, de médianes ou d'autres mesures de tendance centrale de quantités reposant sur certains ou sur tous les membres de la population visée par l'enquête. Elles comprennent aussi spécifiquement des estimations de la forme \hat{X} / \hat{Y} où \hat{X} est une estimation de la quantité totale pour la population visée par l'enquête et \hat{Y} , est une estimation du nombre de personnes dans la population visée par l'enquête qui contribuent à cette quantité totale.

Un exemple d'estimation quantitative est, pour les enfants nés avant la date prévue de la naissance, le nombre moyen de semaines avant la date prévue de naissance que ces enfants sont nés. Le numérateur est une estimation du nombre total de semaines avant la date prévue de la naissance où les enfants sont nés et son dénominateur est le nombre d'enfants qui sont nés avant la date prévue de la naissance.

Exemples de questions quantitatives :

Q: Combien de semaines avant ou après terme est-ce que (l'enfant) est né?
R: |_|_| Semaines

Q: Quel était son poids à la naissance, en kilogrammes et en grammes ou en livres et en onces?
R: |_|_|_|_|_|_| Kilogrammes et grammes / |_|_| Livres |_|_| Onces

10.1.2 Totalisation d'estimations catégoriques

On peut obtenir des estimations du nombre d'enfants possédant une certaine caractéristique à partir du fichier de microdonnées en additionnant les poids finaux de tous les enregistrements possédant la ou les caractéristiques qui nous intéressent. On obtient des proportions et des rapports de la forme \hat{X} / \hat{Y} en :

- a) additionnant les poids finaux des enregistrements présentant la caractéristique qui nous intéresse pour le numérateur (\hat{X}),
- b) additionnant les poids finaux des enregistrements présentant la caractéristique qui nous intéresse pour le dénominateur (\hat{Y}), puis en
- c) divisant l'estimation a) par celle de b) (\hat{X} / \hat{Y}).

10.1.3 Totalisation d'estimations quantitatives

On peut obtenir des estimations de quantités à partir du fichier de microdonnées en multipliant la valeur de la variable qui nous intéresse par le poids final de chaque enregistrement, puis en additionnant cette quantité pour tous les enregistrements qui nous intéressent. Pour obtenir, par exemple, une estimation du nombre total de semaines avant la date prévue de la naissance où les enfants sont nés, multipliez la valeur déclarée à la question GHLCd4DB (nombre de semaines pour lesquelles l'enfant est né avant la date prévue de la naissance) par le poids final de l'enregistrement, puis additionnez cette valeur pour tous les enregistrements où la variable GHLCd4DA = 1 (l'enfant est né avant la date prévue de la naissance).

Pour obtenir une moyenne pondérée de la forme \hat{X} / \hat{Y} , le numérateur (\hat{X}) est calculé comme une estimation quantitative et le dénominateur (\hat{Y}) est calculé comme une estimation catégorique. Pour estimer, par exemple, le nombre moyen de semaines avant la date prévue de la naissance où les enfants sont nés,

- a) estimez le nombre total de jours (\hat{X}) tel qu'il est décrit ci-dessus,
- b) estimez le nombre d'enfants (\hat{Y}) inclus dans cette catégorie en additionnant les poids finaux de tous les enregistrements où la variable GHLCd4DA = 1, puis
- c) divisez l'estimation a) par l'estimation b) (\hat{X} / \hat{Y}).

10.2 Lignes directrices pour la diffusion de coefficients de variation

Avant de diffuser et/ou de publier toute estimation établie à partir de l'Enquête sur les communautés, les utilisateurs devraient premièrement déterminer le niveau de qualité de cette estimation. Les niveaux de qualité sont *acceptable*, *médiocre* et *inacceptable*. Les erreurs d'échantillonnage et non dues à l'échantillonnage influencent la qualité des données. Cependant, pour les communautés avec un plan d'échantillonnage, on ne déterminera le niveau de qualité d'une estimation qu'à partir d'une erreur d'échantillonnage dont rend compte le coefficient de variation indiqué à l'intérieur du tableau qui figure ci-dessous.

On devrait premièrement déterminer le nombre d'enfants retenus pour le calcul de l'estimation. En général, si ce nombre est inférieur à 30, il faudrait considérer l'estimation pondérée comme étant de qualité inacceptable. Pour les communautés, en particulier si la fraction d'échantillonnage est grande ou si un recensement a été effectué, le nombre minimal pourrait être beaucoup moindre. Cependant, l'information diffusée ou tabulaire devra satisfaire les règles de confidentialité en ce qui a trait aux comptes minimaux d'échantillon qui pourront être diffusés.

Pour les estimations pondérées basées sur les tailles d'échantillon de 30 ou plus, les utilisateurs devraient déterminer le coefficient de variation de l'estimation et suivre les lignes directrices relatives au niveau de qualité qui figurent ci-dessous. Celles-ci devraient être appliquées, aux estimations pondérées arrondies.

On peut considérer qu'il est possible de divulguer toutes les estimations. Cependant, celles d'un niveau de qualité médiocre ou inacceptable doivent être accompagnées d'une mise en garde pour avertir les utilisateurs subséquents, et ne doivent pas aller à l'encontre des règles de confidentialité.

Une attention spéciale devra être donnée aux estimations provenant de communautés où un recensement de tous les enfants a été effectué. Même si l'erreur due à l'échantillonnage n'est pas applicable aux estimations produites, les erreurs non dues à l'échantillonnage peuvent prendre une importance plus significative. Les poids contenus sur le fichier de ces communautés sont produits pour corriger le sous-dénombrement en raison du nombre de ménages non-répondants. Il y a d'autres raisons liées aux erreurs non dues à l'échantillonnage qui expliquent le fait d'avoir un estimateur imparfait. Même si un échantillon ou un recensement dans le cadre d'une enquête peuvent tous deux produire ce type d'erreur, un estimateur basé sur un recensement est souvent plus vulnérable puisqu'il devrait produire une information exacte.

Lignes directrices relatives au niveau de qualité de l'estimation

Niveau de qualité de l'estimation	Lignes directrices
1) Acceptable	<p>Les estimations proviennent d'une taille d'échantillon de 30 ou plus, et présentent de faibles coefficients de variation, de l'ordre de 0,0 à 16,5 %.</p> <p>Aucune mise en garde n'est requise.</p>
2) Médiocre	<p>Les estimations proviennent d'une taille d'échantillon de 30 ou plus, et présentent des coefficients de variation élevés, de l'ordre de 16,6 à 33,3 %.</p> <p>Ces estimations devraient être signalées par la lettre M (ou un quelconque identificateur similaire). Elles devraient être accompagnées d'une mise en garde avertissant les utilisateurs subséquents des niveaux élevés d'erreurs associés aux estimations.</p>
3) Inacceptable	<p>Les estimations proviennent d'une taille d'échantillon inférieure à 30, ou présentent des coefficients de variation très élevés, supérieurs à 33,3 %.</p> <p>Statistique Canada recommande de ne pas diffuser d'estimations de qualité inacceptable. Si un utilisateur choisit cependant de le faire, ces estimations devraient alors être signalées à l'aide de la lettre I (ou d'un quelconque identificateur similaire) et devraient être accompagnées de la mise en garde suivante :</p> <p>« Nous informons l'utilisateur que ces estimations (désignées avec la lettre I) ne respectent pas les normes de qualité de Statistique Canada. Les conclusions qui reposeront sur ces données ne seront pas fiables et seront très probablement invalides. »</p>

11.0 Variabilité d'échantillonnage

L'échantillon de l'Enquête sur les communautés de 2005 a été prélevé selon la méthode d'échantillonnage aléatoire simple (EAS). Les coefficients de variation (CV) applicables à une grande variété d'estimations catégoriques produites à partir de ce fichier de microdonnées peuvent facilement être calculés par l'utilisateur à l'aide de fonctions types (formules) offertes par la plupart des progiciels statistiques ou tableurs.

Les coefficients de variation doivent être calculés à l'aide de la formule de la variance pour un échantillonnage aléatoire simple et du facteur de correction d'échantillonnage pour population finie (fpc). Avant d'utiliser un logiciel statistique ou tableur, il convient de s'assurer que les formules appropriées sont utilisées. Le tableau suivant présente diverses expressions algébriques pour la variance, l'erreur-type, le coefficient de variation et l'intervalle de confiance pour un échantillon aléatoire simple.

Estimation	Variance	Erreur-type	Coefficient de variation	Intervalle de confiance (1- α) %
Moyenne (\bar{X})	$(1-f) S^2 / n$	$((1-f)/n)^{1/2} S$	$((1-f)/n)^{1/2} S / \bar{X}$	$\bar{X} \pm t_{\alpha/2} ((1-f)/n)^{1/2} S$
Total (\hat{N})	$N^2 (1-f) S^2 / n$	$N((1-f)/n)^{1/2} S$	$((1-f)/n)^{1/2} S / \hat{N}$	$\hat{N} \pm t_{\alpha/2} N((1-f)/n)^{1/2} S$
Proportion (p)	$\frac{(1-f) PQ N}{(N-1)(n-1)}$	$\frac{((1-f) PQ N)^{1/2}}{((N-1)(n-1))^{1/2}}$	$\frac{((1-f) Q N)^{1/2}}{(P(N-1)(n-1))^{1/2}}$	$\frac{p \pm t_{\alpha/2} ((1-f) PQ N)^{1/2}}{((N-1)(n-1))^{1/2}}$

Notation

N : nombre d'unités dans la population

n : nombre d'unités dans l'échantillon

$f = n/N$: taux d'échantillonnage de la population

$fpc = 1 - f$: facteur de correction d'échantillonnage pour population finie

y_i : valeur de la caractéristique y pour l'unité i^{th}

$P = N_y / N$: proportion d'unités de la population ayant l'attribut y

$Q = \frac{1-P}{N}$: proportion d'unités de la population n'ayant pas l'attribut y

$S^2 = \sum_{i=1} (y_i - \bar{y})^2 / (N-1)$: variance de la caractéristique y dans la population

Comme il est mentionné dans le chapitre précédent, la fréquence par cellule acceptable pour l'échantillon de l'Enquête sur les communautés est de 30 unités d'échantillonnage répondantes (ou un chiffre de population estimé de 164). Le tableau suivant montre la diminution de précision qui se produit lorsque la fréquence par cellule baisse.

Taille de l'échantillon	Taille estimée de la population	Coefficient de variation
29	159	16,47 %
28	153	16,78 %
8	44	31,80 %
7	38	34,02 %

12.0 Pondération

Le principe qui sous-tend une estimation pour un échantillon probabiliste veut que chacune des personnes incluses dans l'échantillon « représente », en plus d'elle-même, plusieurs autres personnes qui en sont exclues. L'étape de la pondération vise à calculer le nombre d'individus de la population qui sont représentés par un enregistrement.

Pour chacune des communautés, tous les répondants se sont vus attribuer des poids égaux, puisque l'échantillonnage pour chaque communauté a été effectué de façon aléatoire, avec probabilités de sélection égales. Le poids attribué à chaque répondant correspond simplement au rapport de la taille de la population à la taille de l'échantillon. Cette valeur a été ajustée pour tenir compte des individus que l'on n'a pu contacter au moment de la collecte et des non répondants.

Pour chaque communauté, si l'on somme les poids des répondants on obtient l'estimation de la taille de la population réelle de la communauté. Par conséquent, cette dernière estimation représente le nombre d'enfants inscrits en maternelle pour cette communauté en septembre 2004 et qui étaient toujours inscrits pendant la période de la collecte.