

Enquête : description et méthodologie



1. Aperçu du devis de l'étude

L'enquête longitudinale décrite en détail dans les pages qui suivent en est une annuelle et représente, lors de la première année de collecte, tous les enfants d'environ 5 mois issus de naissances simples et nés au Québec. La première collecte a été fixée au moment où la majorité des nourrissons avaient environ 5 mois, car c'est l'âge auquel la plupart des nourrissons sont suffisamment développés pour que plusieurs données centrales aux fins de l'étude puissent être colligées. Chaque collecte annuelle se déroule sur une période de 6 mois et chaque mois de collecte est appelé « vague de collecte » ; les trois premières vagues de collecte se déroulent entre mars et juin et les trois dernières entre septembre et décembre. Ces vagues sont généralement disjointes et chacune d'entre elles se déroule au cours d'une période qui dure de 4 à 5 semaines. Ainsi, à part de très rares exceptions, les enfants sont tous rencontrés au moment où ils ont un âge comparable. Cette « période d'admissibilité » de l'enfant est très importante, compte tenu de la rapidité du développement chez les tout jeunes enfants. Par la suite, un délai de 12 mois sépare les collectes annuelles. Afin d'avoir des données valables et fiables, on a opté, comme on le verra tout au long de ce numéro, pour un échantillon d'enfants relativement important et réparti à travers le Québec. Ce faisant, des contraintes reliées au temps de collecte, au traitement des données annuelles, à la création des instruments et aux coûts d'une entreprise de cette envergure ont orienté certaines décisions, entre autres celle de concentrer la collecte des données annuelles sur une période de 6 mois et non de l'échelonner tout au long de l'année de même que le fait de rencontrer les familles une seule fois par année pour une période de 5 ans. Dans l'ÉLDEQ 1998-2002³, chaque année d'enquête est appelée « volet ». Ainsi, les données publiées cette année concernent les enfants d'environ 5 mois, donc le « volet 1998 », 1998 désignant l'année au cours de laquelle les enfants étaient âgés de 5 mois. Et si, comme souhaité, l'enquête auprès des enfants de l'ÉLDEQ 1998-2002 se poursuit lorsque ces derniers feront leurs études primaires, nous entrerons dans la deuxième « phase » de l'ÉLDEQ 1998-2002, la première phase (ou phase I) étant exclusivement consacrée aux cinq collectes

annuelles qui auront permis de mieux comprendre le développement au cours de la « petite enfance ».

1.1 État d'avancement des travaux

Le présent rapport, publié sous la forme d'une collection de numéros, porte sur des données transversales recueillies auprès d'un large échantillon de nourrissons âgés d'environ 5 mois en 1998. Il rend compte de la première d'une série de cinq mesures annuelles auprès de 2 120 enfants⁴ du Québec qui seront suivis jusqu'à l'âge de 5 ans. L'*Étude longitudinale du développement des enfants du Québec* (ÉLDEQ 1998-2002) a débuté en 1996 et se terminera par la publication du troisième et dernier volume en 2004. Au moment de la publication du volume 1, les collectes de quatre des cinq prétests prévus ont été effectuées auprès des enfants alors qu'ils étaient âgés de 5 mois, de 17 mois, de 29 mois et de 41 mois et les collectes des deux premiers temps de mesure de l'enquête proprement dite ont été réalisées (enfants âgés de 5 et de 17 mois) (voir tableau 1.1). En 2000, les enfants de 29 mois et leur famille sont interviewés et la préparation du dernier prétest auprès des enfants de 53 mois se termine. Les deuxième et troisième collections de numéros de l'ÉLDEQ 1998-2002, soit les volumes 2 et 3, devraient être publiées respectivement en 2002 et en 2004.

3. L'enquête est aussi appelée « En 2002... J'aurai 5 ans! » sur tous les outils de communication que nous avons développés pour les parents et les partenaires de la Direction Santé Québec. De plus, les années 1998-2002 qui accompagnent l'acronyme ÉLDEQ identifient les années de collecte de ce qui pourrait constituer la première d'une série de phases de l'ÉLDEQ.

4. Pour la première année de collecte, les résultats concernant 2 223 enfants ont été conservés. Pour le suivi longitudinal, seuls 2 120 enfants seront inclus, les 103 retranchés étant une partie d'un suréchantillon servant à mesurer les effets de la tempête de verglas de janvier 1998.

Tableau 1.1

Récapitulatif des collectes de prétests et d'enquêtes et des publications, selon l'année, 1998

Année	Prétest	Enquête	Publication
1996	Nourrissons de 5 ms
1997	Enfants de 17 ms
1998	Enfants de 29 ms	Nourrissons de 5 ms	Rapport préliminaire ⁵ Prétest 5 ms
1999	Enfants de 41 ms	Enfants de 17 ms	...
2000	Enfants de 53 ms	Enfants de 29 ms	Rapport transversal (Vol. 1) Enquête 5 ms
2001	...	Enfants de 41 ms	...
2002	...	Enfants de 53 ms	1 ^{er} rapport longitudinal (Vol. 2) 5, 17, 29 ms
2003
2004	2 ^e rapport longitudinal (Vol. 3) 5 à 53 ms

5. SANTÉ QUÉBEC, JETTÉ M., H. DESROSIERS et R.E. TREMBLAY (1997).

2. Présentation générale de l'enquête de 1998 auprès des enfants de 5 mois (ÉLDEQ 1998-2002)

Le premier numéro de ce rapport décrit de manière détaillée les différents aspects méthodologiques de l'enquête « En 2002... J'aurai 5 ans! ». Il constitue un OUTIL INDISPENSABLE à une lecture éclairée des données d'enquête présentées dans tous les autres numéros de cette première collection. Les informations qui suivent devraient permettre au lecteur d'interpréter les résultats en les replaçant dans leur contexte de production. Seront donc abordés : 1) les instruments d'enquête, leur fonction, leur contenu et les aspects logistiques qui les régissent; 2) la collecte des données et les taux de collaboration aux principaux instruments de l'enquête; 3) les aspects statistiques importants tels la base de sondage, le plan de sondage, la taille de l'échantillon, les taux de réponse, etc., autant de normes auxquelles est assujettie la valeur des résultats de l'enquête et qui mettent en lumière la portée et les limites des données de ce premier volet.

2.1 Les instruments d'enquête et de collecte

La première phase de cette enquête longitudinale s'étalant sur 5 ans s'articule autour de 14 instruments permettant soit la collecte des données, soit le suivi longitudinal des familles. Puisque l'objectif de ce premier numéro est de faciliter la bonne compréhension de l'enquête dans son ensemble, les informations plus détaillées relatives à la source et à la justification des questions, à la construction des variables dérivées, d'échelles et d'indices ainsi qu'aux modifications apportées aux questions originales sont consignées dans un autre numéro de cette collection, le dernier⁶, prévu pour décembre 2000. Quant aux instruments d'enquête proprement dits, ils sont disponibles en tout temps à la Direction Santé Québec.

Afin de mieux saisir la finalité de chacun des instruments qui existent, tant en versions française qu'anglaise, ils seront présentés suivant l'ordre d'introduction dans le processus de collecte.

Instruments #1 et #2 : Lettre d'introduction et brochure de présentation de l'enquête

À partir du *Fichier maître des naissances* du ministère de la Santé et des Services sociaux (MSSS), des enfants ont été sélectionnés. Comme les différents aspects touchant la base de sondage, l'échantillon, la collecte, etc. seront abordés plus loin, soulignons seulement ici que c'est à partir des adresses contenues dans le *Fichier maître des naissances* que la lettre d'introduction accompagnée de la brochure explicative ont été acheminées aux familles visées par l'enquête.

Cette LETTRE et la BROCHURE parvenaient aux familles des enfants sélectionnés au moins 2 semaines avant qu'elles soient contactées par la maison de sondages engagée à cet effet. Ces deux instruments ont plusieurs fonctions : présenter les objectifs généraux et spécifiques de l'enquête; présenter l'organisme Santé Québec, son partenariat avec le MSSS, le réseau de la santé et des services sociaux et les chercheurs associés au projet; expliquer le rôle important des parents; présenter les particularités d'une enquête longitudinale; mentionner explicitement la durée des entrevues et l'incitatif symbolique remis aux parents; rassurer les parents quant à la stricte confidentialité des données recueillies; signaler que les adresses ont été obtenues du MSSS avec l'accord préalable de la Commission d'accès à l'information du Québec (CAI); et, finalement, annoncer aux parents l'appel téléphonique du Bureau d'interviewers professionnels (maison de sondages BIP) pour les convier à une entrevue en face à face (c'est-à-dire de personne à personne).

Instrument #3 : Pochette-souvenir « En 2002... J'aurai 5 ans! »

Une pochette-souvenir avec l'emblème de l'enquête est remise aux parents lors de la première des cinq visites annuelles afin qu'ils puissent conserver en un même endroit les documents qu'ils reçoivent périodiquement de Santé Québec et du BIP.

6. Ce numéro présentera, notamment, la/les sources et justifications de chaque question et échelle utilisées dans l'enquête auprès des enfants de 5 mois. De même, il rendra compte des instruments qui ont été mis au point spécifiquement pour ce premier volet de l'enquête de cohorte.

Instrument #4 : Formulaire de consentement libre et éclairé

Ce formulaire, soumis et accepté par le Comité d'éthique de la Direction Santé Québec de l'ISQ, devrait être signé annuellement par l'un ou l'autre ou les deux parents de l'enfant. Ce formulaire n'est pas obligatoire pour que le ou les parents participent à l'enquête mais tous les ans les parents sont conviés à le signer. Il précise le partenariat avec les chercheurs, confirme que le contenu de l'enquête a bien été expliqué aux parents, réitère les aspects de confidentialité et d'anonymat qui régissent l'enquête et assure aux parents le droit de refuser ou de se retirer de l'enquête sans préjudice et en tout temps. Fournis en deux exemplaires par ménage, les formulaires sont signés et datés par les parents et contresignés par l'intervieweuse⁷. Les parents en conservent un original qu'ils rangent dans la pochette et l'intervieweuse rapporte le second à la maison de sondages.

Instrument #5 : Questionnaire informatisé rempli par l'intervieweuse (QIRI)

Instrument maître de l'ÉLDEQ 1998-2002, il provient presque intégralement de l'*Enquête longitudinale nationale sur les enfants et les jeunes* (ELNEJ) conduite par Statistique Canada depuis 1994 auprès de plus de 20 000 enfants originellement âgés de 0 à 11 ans. Il permet la collecte d'informations concernant les membres du ménage dont l'enfant cible de 5 mois fait partie. Il est administré par l'intervieweuse et les réponses aux questions sont obtenues de la personne qui connaît le mieux l'enfant (PCM) : dans la quasi-totalité des cas, la mère biologique de l'enfant. Ce questionnaire contient annuellement quelque 600 variables et requiert, en moyenne, entre 50 et 60 minutes pour être rempli. Les thèmes abordés sont nombreux et variés : renseignements sociodémo-graphiques et socioéconomiques sur les membres du ménage; conditions de logement; habitudes de vie et état de santé des parents (et de la mère pendant la grossesse); tempérament et développement moteur et social du nourrisson; fonctionnement de la famille; rôle parental; antécédents familiaux et garde légale des enfants; et, finalement, recours aux services de garde et sécurité du

quartier. Chacun des thèmes de l'enquête et sa localisation dans les principaux instruments de collecte est présenté au tableau 2.1 en fin de section 2.1.

L'une des particularités de l'ÉLDEQ 1998-2002 est qu'elle s'intéresse aux deux parents biologiques, qu'ils soient présents ou non dans le ménage. Dans le QIRI, si le père biologique est absent (PBA) ou si la mère biologique est absente (MBA), un court questionnaire est utilisé afin que des informations soient recueillies en ce qui a trait à son âge, ses études, son travail et la fréquence de ses contacts avec l'enfant cible. De plus, l'intervieweuse demande à la PCM de lui transmettre les coordonnées du parent biologique absent afin que la maison de sondages lui fasse parvenir un questionnaire autoadministré à retourner par la poste.

Enfin, une section du QIRI est consacrée aux personnes de référence. Ces questions permettent d'obtenir les coordonnées, les plus précises possible, d'au moins deux personnes que la PCM aviserait d'un éventuel changement d'adresse (ou de numéro de téléphone). Ainsi, lorsque la maison de sondages ne parvient pas à rejoindre un ménage lors d'une année donnée, elle peut prendre contact avec l'une ou l'autre de ces personnes pour obtenir les nouvelles coordonnées de la famille ou encore pour connaître la raison pour laquelle il est impossible de la joindre.

Instrument #6 : Questionnaire papier rempli par l'intervieweuse (QPRI)

Il vient compléter le QIRI, donc l'entrevue en face à face avec la PCM. Les 72 variables mesurées dans ce questionnaire se regroupent autour des thèmes suivants : information concernant les familles maternelle et paternelle de l'enfant cible; perception de la situation socioéconomique; mode d'alimentation du nourrisson; attitudes de l'entourage à l'égard de l'allaitement; prise de vitamines et de minéraux; et habitudes de vie reliées à la santé buccodentaire.

Instrument #7 : Formulaire d'autorisation de communiquer des renseignements contenus aux dossiers de la mère et du bébé

Ce formulaire doit obligatoirement être signé par la mère biologique du nourrisson pour que la Direction Santé Québec puisse l'utiliser. De plus, il est contresigné par l'intervieweuse, agissant à titre de témoin, et a une durée

7. Tous les intervieweurs de cette enquête étant de sexe féminin, nous utiliserons, dans la suite du texte, le terme intervieweuse pour les désigner.

de vie légale limitée à 90 jours entrant en vigueur à la date de signature de la mère. Il permet de recevoir des hôpitaux ou des maisons de naissance où les mères ont accouché les renseignements suivants : résumé du dossier obstétrical complet de la mère; compte rendu de l'examen anatomopathologique du placenta et la « Feuille sommaire du CH de courte durée » de la mère; résumé du dossier complet de l'enfant cible incluant la « Feuille sommaire du CH de courte durée », l'examen objectif du nouveau-né et le compte rendu de l'examen du sang au cordon ombilical. Environ 50 variables concernant la mère et l'enfant ont été identifiées dans un premier temps (ex. : type d'accouchement, complications périnatales, recours à des soins spécialisés pour le nourrisson, etc.) et font l'objet d'un numéro à paraître en 2001.

Instrument #8 : Jeu 1, 2, 3 mains

Cette tâche de type « Imitation de placement d'objets (IPO) », mise au point initialement pour des enfants âgés de 10 et 12 mois, a dû être modifiée pour s'adapter aux capacités d'un enfant de 5 mois. Ce nouvel instrument permet une première mesure de la capacité mentale attentionnelle du nourrisson et propose deux niveaux de difficulté. Au premier niveau, l'enfant doit coordonner les schèmes de vision et de préhension afin de regarder un objet qu'on lui montre et de saisir cet objet en ayant, au préalable, ouvert les mains. Au deuxième niveau de difficulté, on augmente le nombre d'éléments que l'enfant doit prendre en compte. En fait, on place un objet dans chaque main du nourrisson et on lui en présente un troisième. L'enfant doit laisser tomber un ou les deux objets pour prendre le troisième. Cette activité est la première mesure « objective » réalisée directement avec le nourrisson et elle exige de lui environ 5 minutes d'attention.

Instrument #9 : Questionnaire autoadministré de la mère (QAAM)

Comme son nom l'indique, ce questionnaire est rempli par la mère, généralement sans aide, à moins qu'elle n'en demande à l'intervieweuse pour des raisons de limitations quant à la lecture, à l'écriture, etc. Cet instrument comprend 88 variables et permet la mesure d'informations plus sensibles tels les grossesses antérieures, certaines difficultés rencontrées à l'adolescence ou lors de la vie adulte (ex. : faire l'école buissonnière, fuguer du domicile familial, etc.) ainsi que le soutien apporté par le conjoint.

Il permet également la collecte d'informations sur les relations mère/enfant et le sommeil du nourrisson. Finalement, en l'absence du père biologique, on propose à la mère de remplir une section supplémentaire concernant la fréquence des visites du père, le soutien financier fourni et certaines difficultés que celui-ci aurait pu rencontrer à l'adolescence ou lors de sa vie adulte (les mêmes que pour la mère). À l'instar des variables figurant dans le QPRI, les questions du QAAM sont, soit empruntées à d'autres enquêtes, soit mises au point aux fins de l'ÉLDEQ 1998-2002, et ne figurent pas dans l'ELNEJ. Le temps requis pour remplir ce questionnaire, comme celui du QAAP décrit ci-après, est d'environ 20 minutes.

Instrument #10 : Questionnaire autoadministré du père (père présent-QAAP; père biologique absent-QAAPABS)

Tout comme le QAAM, le QAAP est généralement rempli sans intervention de l'intervieweuse, à moins que le père ne sollicite son aide. Constituant le pendant du QAAM et d'une partie du QIRI, il contient 74 variables concernant le tempérament de l'enfant cible, les relations père/enfant, une section sur la santé mentale et l'une concernant certaines difficultés que le père aurait pu éprouver lors de son adolescence ou de sa vie adulte.

Une des caractéristiques de cet instrument est, qu'en plus d'être un questionnaire original à l'ÉLDEQ 1998-2002, il peut être rempli tant par le père biologique absent du ménage(QAAPABS) que par le nouveau conjoint de la mère. Cette même particularité s'applique au QAAM en regard de la mère biologique et de la nouvelle conjointe du père. À un même nourrisson peuvent donc correspondre un QAAP (ou QAAM) et un QAAPABS (ou QAAMABS). Ainsi, il sera peut-être possible de conserver, comme informateur pertinent, le parent biologique absent malgré l'arrivée d'une nouvelle figure masculine ou féminine dans la vie de l'enfant cible. À noter cependant que seuls les parents biologiques absents ayant un contact avec l'enfant au moins une fois par mois, et dont on a obtenu les coordonnées, sont admissibles à recevoir leur QAAMABS ou QAAPABS.

Instrument #11 : Journal de bébé

Sous la forme d'un agenda précis aux 5 minutes, ce journal qui couvre une période d'observation de 48 heures peut être rempli par la PCM ou toute personne qui donne des soins ou s'occupe de l'enfant cible. Il a comme

objectif de mesurer la fréquence et la durée de certains comportements de l'enfant tels que dormir, être éveillé et content, éveillé et maussade, éveillé et en pleurs, pleurer de manière inconsolable et boire (au sein ou au biberon). De plus, il permet de consigner la fréquence et le temps que lui consacrent les adultes qui en ont la garde en ce qui a trait aux soins de base (baigner, changer de couche, habiller) et aux contacts physiques (c'est-à-dire lorsque l'enfant est tenu dans les bras d'un adulte). Le continuum de temps est imagé par des règles dont les unités sont des périodes de 5 minutes. De plus, chacune des règles d'une durée de 6 heures (4 règles par jour d'observation doivent être complétées, donc 8 règles au total) est divisée en deux pour que la partie supérieure de la règle permette de dessiner les comportements de l'enfant et la partie inférieure, ceux des adultes. Ainsi, les analyses des comportements peuvent être faites séparément pour l'enfant et les adultes mais également en concomitance (voir spécimen à l'annexe 1). Le temps requis pour remplir ce questionnaire est, en moyenne, 30 minutes par jour d'observation.

Instrument #12: Inventaire du milieu familial (IMF)

Seul questionnaire rempli exclusivement par l'intervieweuse sans la présence des parents, il permet, par le biais d'environ 40 variables, de consigner des observations faites en cours d'entrevue en regard des interactions mère/enfant, de l'organisation de l'environnement physique de l'enfant et de certains comportements de la mère et du bébé.

Instrument #13 : Questionnaire sur la tempête de pluie verglaçante (janvier 1998)

Ce questionnaire a été ajouté *in extremis* aux instruments de l'ÉLDEQ 1998-2002 afin de déterminer l'impact de la tempête de pluie verglaçante de janvier 1998 et ses conséquences sur les estimateurs de l'enquête. Court instrument d'une dizaine de questions, rempli en face à face avec la PCM dans tous les ménages de l'enquête, il a permis d'évaluer la proportion de l'échantillon qui avait été touchée par ce phénomène naturel et d'estimer les divers degrés d'atteinte.

Après avoir recueilli les données, des croisements entre diverses variables ont été effectués : certaines variables qui vraisemblablement auraient pu être modifiées par l'épisode de verglas et ses suites et d'autres variables

ne devant pas, à première vue, être influencées par ce caprice climatique. Les analyses préliminaires n'ont révélé aucun biais potentiel attribuable à cet épisode climatique. Ainsi, l'éventuel stress vécu par les enfants et les parents ne semble pas avoir eu d'effets directs ou indirects détectables sur les dimensions mesurées du développement des enfants, et ce, quel qu'ait été l'âge de l'enfant (nourrisson dans le 2^e trimestre de gestation (grossesse), 3^e trimestre ou 3 premiers mois de vie⁸). Compte tenu de l'absence de résultat significatif, aucun résultat produit à partir de ce questionnaire ne sera publié dans le cadre de la présente collection. Même si ces résultats sont très rassurants pour la qualité des données produites dans la première année de la cohorte, ils demeurent disponibles s'il advenait que des analyses longitudinales soulèvent des interrogations.

Instrument #14 : Carte d'anniversaire à l'enfant cible

Une carte d'anniversaire est envoyée par la Direction Santé Québec au ménage dans les jours entourant chaque anniversaire de l'enfant. Santé Québec y insère également un carton de rappel concernant la mise à jour, si nécessaire, des coordonnées du ménage (adresse, numéro de téléphone, etc.).

Terminons cette section relative à la description des instruments de l'enquête en mentionnant que les familles reçoivent semestriellement des informations concernant la cohorte sous la forme d'un *Communiqué aux parents* de l'enquête « En 2002... J'aurai 5 ans! ». En plus de garder les parents informés de l'état d'avancement de l'enquête, de fournir des réponses à des questions récurrentes qu'ils nous transmettent par l'intermédiaire des sections « commentaires » mises à leur disposition dans le QAAM, le QAAP et le Journal de bébé, ces communiqués sont autant d'occasions de rappeler aux parents l'importance de participer à l'enquête et d'aviser d'un changement d'adresse, si nécessaire.

8. On verra ultérieurement (section 3.1.3.) qu'en janvier 1998 les nourrissons de la cohorte n'avaient pas tous le même âge étant donné que la collecte annuelle de données s'échelonne sur 6 mois, soit de la mi-mars à la mi-juin et du début de septembre au début de décembre d'une même année.

Tableau 2.1

Récapitulation des thèmes abordés dans l'enquête auprès des nourrissons de 5 mois (ÉLDEQ 1998-2002) selon l'instrument de collecte, 1998

Thèmes de l'ÉLDEQ 1998-2002 année 1998 : nourrissons de 5 mois	QIRI Inst. 5 ^a	QPRI Inst. 6	QAA M Inst. 9	QAAP ^b Inst. 10	Dos. méd. Inst.7	J. de bébé Inst. 11	Jeu 1,2,3 mains Inst. 8	IMF Inst. 12
PARENTS (MÈRE ET PÈRE/CONJOINT)								
Renseignements sociodémographiques	√							
Scolarité	√							
Occupation/travail	√							
Santé physique(état gén. et m. chronique)	√							
Habitudes de vie (tabac, alcool, drogues)	√(PCM)			√				
Santé mentale (dépression)			√	√				
Certains comportements antisociaux			√	√				
Soutien apporté par le conjoint			√	√				
Loisirs								
MÈRE BIOLOGIQUE								
Habitudes de vie (durant la grossesse) :	√							
tabac	√							
alcool	√							
médicaments	√							
drogues	√							
Dépression postnatale								
Hospitalisation et soins médicaux spéciaux (après accouchement)	√		√		√			
Grossesses/fécondité	√							
Travail après la naissance					√			
Autres renseignements périnataux								
PARENT BIOLOGIQUE ABSENT								
Fréquence des visites à l'enfant cible et soutien financier	√		√ ^(c)					
Scolarité	√							
Occupation et travail			√ ^c					
Certains comportements antisociaux								

(suite du tableau à la page suivante)

- Réfère aux numéros d'instruments de collecte tels qu'ils ont été présentés à la section 2.1.
- Le QAAP s'adresse aux pères/conjoints présents et aux pères biologiques absents du ménage (voir section 2.1, instrument 10).
- Ces renseignements sur le père biologique absent sont recueillis par procuration auprès de la mère biologique.

Thèmes de l'ÉLDEQ 1998-2002 année 1998 : nourrissons de 5 mois	QIRI Inst. 5 ^a	QPRI Inst. 6	QAAM Inst. 9	QAAP ^b Inst. 10	Dos. méd. Inst.7	J. de bébé Inst. 11	Jeu 1,2,3 mains Inst. 8	IMF Inst. 12
ENFANT CIBLE - 5 MOIS								
Renseignements sociodémographiques	√							
Santé physique :	√							
état général	√							
poids et taille	√							
poids à la naissance	√							
blessures	√							
maladies chroniques	√							
consultations médicales	√							
hospitalisation					√			
autres renseignements périnataux		√						
santé buccodentaire								
Alimentation :		√						
mode d'alimentation		√						
vitamines et minéraux		√						
introduction d'aliments solides		√						
attitudes à l'égard de l'allaitement		√	√					
Sommeil	√			√				
Tempérament						√		√
Comportements du nourrisson	√							
Développement moteur et social							√	
Développement cognitif	√							
Activités	√							
Modalités de garde								
RELATIONS PARENTS/ENFANTS	√							√
Pratiques parentales	√							
Lecture faite à l'enfant						√		
Soins prodigués/contacts physiques avec l'enfant			√					√
Relations mère/enfant				√				
Relations père/enfant								
FAMILLE ET MÉNAGE	√							
Membres constituant le ménage		√						
Informations intergénérationnelles (familles maternelle et paternelle)	√							
Antécédents familiaux et garde légale de l'enfant	√							
Fonctionnement de la famille	√	√						
Revenu du ménage		√						
Perception de la situation socioéconomique	√							√
Environnement physique :	√							
conditions de logement								
sécurité du quartier								

Source : Institut de la statistique du Québec, ÉLDEQ 1998-2002.

2.2 La collecte de données

Après le tirage de l'échantillon de chacune des vagues de collecte⁹ par la Direction de la méthodologie et des enquêtes spéciales de l'ISQ (voir les sections 3 et suivantes), les données étaient acheminées au Service du support aux opérations de la Régie de l'assurance-maladie du Québec (RAMQ). À cette étape, un jumelage des données¹⁰ concernant les enfants sélectionnés était fait entre les données des deux fichiers administratifs, soit celui des naissances et celui de la RAMQ. Cette opération servait à mettre à jour les adresses figurant au *Fichier maître des naissances* (données qui dataient d'environ 5 mois) grâce à celles intégrées au fichier de la RAMQ qui étaient plus récentes dans environ 3 % des cas. Ainsi, la fiabilité des données de la base de sondage « évolutive » de l'enquête s'en trouvait améliorée. Au total, après le jumelage, la base de sondage est demeurée imprécise pour un peu moins de 6 % de l'échantillon, c'est-à-dire qu'un total de 172 familles sélectionnées n'ont pu être rejointes à cause de coordonnées inexactes (adresse civique et numéro de téléphone).

À la suite de l'opération de jumelage, la RAMQ faisait parvenir à la Direction Santé Québec la banque de données de chacune des six vagues de collecte et les envois postaux des lettres d'introduction et des brochures étaient faits à partir de ces coordonnées. Conséquemment à chacune des six opérations postales, la maison de sondages recevait la liste des adresses et des numéros de téléphone des répondants potentiels pour chacune des vagues. Au total, 2 940 ménages ont été sélectionnés (voir tableau 2.2) et ont fait l'objet de plusieurs tentatives pour être rejoints. Si le contact téléphonique ne pouvait être établi à partir des coordonnées disponibles, la recherche dans les bottins téléphoniques et les visites aux domiciles des personnes non localisées constituaient les étapes suivantes. Lorsque le numéro de téléphone demeurait introuvable et que

personne ne répondait à la porte du domicile identifié, les voisins immédiats du domicile étaient sollicités afin d'aider à la localisation des familles recherchées. Ce n'est qu'après toutes ces démarches que les ménages et/ou les enfants étaient considérés introuvables.

Tableau 2.2
Échantillon initial de l'ÉLDEQ 1998-2002 et taux de collaboration, 1998

	Bébé cible	
	n	%
Échantillon de départ	2 940	100,0
Familles introuvables (coordonnées erronées)	172	5,9
Familles exclues	93	3,2
Familles et bébés possibles à contacter pour l'enquête	2 675	90,1
Familles et bébés non rejoints (après plusieurs tentatives)	14	0,5
Familles ayant refusé	438	16,4
Total refus + non rejoints	452	16,9
Taux de collaboration	2 223^a	83,1

a. Rappelons que des 2 223 nourrissons qui ont fait l'objet de l'enquête en 1998, seulement 2 120 seront conservés pour le suivi longitudinal.

Source : Institut de la statistique du Québec, *ÉLDEQ 1998-2002*.

Par ailleurs, même lorsque le contact téléphonique a pu être établi, d'autres enfants et leurs familles, 93 au total, n'ont pu être inclus dans l'étude pour diverses raisons : le décès du bébé sélectionné (5 nourrissons); des familles faisant déjà l'objet d'un suivi longitudinal (5 au total) que la Direction Santé Québec ne voulait pas surexposer (suivi soit par l'ELNEJ, le prétest de l'ÉLDEQ 1998-2002 ou l'*Étude des jumeaux nouveau-nés du Québec* [ÉJNQ]); des familles ne parlant ni français ni anglais (81 familles) pour lesquelles Santé Québec n'a pas adapté d'instruments; et, finalement, 2 nourrissons n'ont pu faire l'objet du suivi car les instruments de l'enquête ne pouvaient mesurer adéquatement le développement d'enfants souffrant de handicaps physiques ou mentaux sévères (tableau 2.2). C'est ainsi que 3,2 % des familles sont exclues de l'enquête, soit en fonction des lignes directrices données par Santé Québec ou des limites des instruments de collecte à mesurer le développement

9. Comme on l'a vu précédemment, est appelée « vague de collecte » une période d'environ 4 semaines à l'intérieur de laquelle les familles des nourrissons doivent être interviewées, et ce, annuellement. La collecte annuelle s'échelonne sur six vagues de collecte, les trois premières vagues au printemps d'une année, en l'occurrence 1998, et les trois dernières à l'automne de la même année.

10. Cette opération de jumelage a dûment et préalablement été autorisée par la Commission d'accès à l'information du Québec (CAI).

d'enfants présentant certaines caractéristiques. Enfin, la maison de sondages a obtenu un taux de collaboration de 83 %, taux pouvant également être interprété comme le taux de collaboration des familles sollicitées (tableau 2.2). Dans les sections consacrées à la méthodologie statistique de l'enquête, les taux de réponse seront présentés et expliqués. Différents des taux de collaboration qui visent essentiellement à rendre compte du déroulement de l'enquête sur le terrain, les taux de réponse servent d'appui à des opérations statistiques élaborées telles que le calcul du poids dû à la non-réponse de chaque individu de l'enquête.

Ainsi, après avoir présenté le nombre total de répondants à l'enquête (taux de collaboration), il est important de connaître le taux de collaboration pour chacun des instruments de collecte proprement dits (tableau 2.3). Il faut mentionner que si le ménage ne remplissait pas le QIRI, aucun des autres instruments ne pouvait être administré. Voilà pourquoi, au niveau de l'enquête, le QIRI représente le dénominateur à partir duquel les taux de collaboration des autres instruments sont calculés.

Tableau 2.3
Liste des instruments de collecte selon le nombre de répondants et les taux de collaboration, 1998

	n	%
QIRI (inst. #5)	2 223	100,0
QPRI (inst. #6)	2 223	100,0
Autorisation pour dossier médical (inst. #7)	2 184	98,2
Jeu 1, 2, 3 mains (inst. #8)	2 120	95,4
QAAM (inst. #9)	2 146	96,5
QAAP (père présent et QAAPABS) (inst. #10)	1 900	85,5
Journal de bébé (inst. #11)	1 782	80,2
IMF (inst. #12)	2 221	99,9

Source : Institut de la statistique du Québec, *ÉLDEQ 1998-2002*.

Compte tenu du nombre élevé d'instruments utilisés dans l'enquête et de leur durée, le taux de collaboration à chaque instrument est particulièrement satisfaisant. En effet, rarissimes sont les enquêtes qui utilisent autant d'instruments dans un devis longitudinal annuel et populationnel. Bien que les taux de collaboration soient bons dans leur ensemble, certains taux, présentés au tableau 2.3, méritent toutefois d'être commentés.

Le « Jeu 1, 2, 3 mains » nécessite la participation d'un enfant de 5 mois. Les règles qui régissent la participation du nourrisson sont sévères et doivent être strictement suivies par les intervieweuses même si cela signifie des taux de collaboration plus faibles. Ainsi, un enfant qui dormait ne devait pas, au nom de la participation à la tâche, être réveillé. De toute manière sa participation aurait été influencée par son état d'endormissement. De même, un enfant souffrant (vaccin, grippe, etc.) n'était pas incité à participer. Malgré cela, 2 120 nourrissons, soit plus de 95 %, ont participé à ce jeu. Par ailleurs, pendant la réalisation du jeu, le bien-être de l'enfant demeurait le souci premier de l'intervieweuse. Cet état de fait peut expliquer en partie que 269 enfants (soit 12,7 %) n'aient pas complété le jeu (données non présentées). Dans le numéro consacré au développement moteur, social et cognitif du nourrisson, ce taux élevé de jeux incomplets fera l'objet d'une discussion plus élaborée. Néanmoins, un fait demeure : en adoptant cette règle d'or, qui est de placer le bien-être du nourrisson avant toute autre considération de recherche ou d'enquête, la Direction Santé Québec assumait l'impact que cette décision aurait sur la quantité et le degré de complétude des données d'enquête recueillies.

En regard du taux de participation des pères au QAAP, il faut mentionner que même s'il est de 10 % plus faible que celui des mères, il demeure relativement élevé compte tenu que les hommes, de manière générale, participent moins aux enquêtes (Groves, 1989; Drew et autres, 1988). À la suite d'un suivi étroit des appels téléphoniques aux ménages, il semble que la présence « protectrice » des pères lors de l'entrevue s'explique, du moins en partie, par le fait que leur plus jeune enfant et la mère de cet enfant étaient sollicités et qu'une intervieweuse, qui leur était inconnue, viendrait à la maison. Il est raisonnable de penser que le fait d'avoir assisté à cette entrevue de 90 minutes a augmenté leur intérêt et, par le fait même, leur niveau de participation. À notre connaissance, il existe peu ou pas d'enquêtes à grande échelle qui ont interpellé des pères, de surcroît, des pères d'aussi jeunes enfants, et comme les autres enquêtes longitudinales de référence ne s'y sont pas intéressées (*Cambridge Study in Delinquent Development*, Angleterre; *Oregon Youth Study* et *National Longitudinal Survey of Youth*, É.-U.; *Christchurch Health and Development Study* et *Dunedin Multidisciplinary Health and Development Study*, Nouvelle-Zélande; *ELNEJ*, Canada), la participation des

pères à l'ÉLDEQ 1998-2002 demeure difficile à situer dans l'univers plus large des enquêtes longitudinales.

Quant au « Journal de bébé », bien que le taux de collaboration apparaît nettement plus faible (80 %), on doit tout de même signaler qu'il est de 6 % supérieur à celui obtenu au prétest des enfants de 5 mois (Santé Québec, Jetté, M. et autres, 1997). Comme mentionné à la section 2.1- Instrument #11, le « Journal de bébé » demande une implication de tous les adultes qui donnent des soins ou qui s'occupent de l'enfant. Cette dernière exigence jumelée à la durée de la période d'observation (48 heures) et au fait que les observateurs devaient remplir cet instrument en reproduisant des dessins à motifs variés, plutôt qu'en recourant à l'écriture, sont autant d'éléments qui ont limité le taux de collaboration.

2.3 Le traitement des données

La validation des données de l'enquête est une étape importante et pratiquement décisive pour le traitement statistique qui sera fait ultérieurement. Dès l'entrée des premiers questionnaires à la maison de sondages, une partie des données doivent être saisies. Elles font l'objet d'une double saisie puis une validation intersaisies est réalisée. En deuxième et dernier lieu, une validation de base des questionnaires papier (rejet de codes inadmissibles, validation des filtres, etc.) relève également de la responsabilité de la maison de sondages qui réalise les collectes. En regard du QIRI, la validation de base est effectuée par la Direction Santé Québec. De la même manière, c'est Santé Québec qui assure, qu'à la troisième étape, les variables sont examinées une à une et croisées avec d'autres variables pour détecter et éliminer ou corriger d'éventuelles valeurs aberrantes. Après avoir validé les données avec des variables complémentaires, Santé Québec réalise une validation interinstruments permettant, par des croisements logiques plus sophistiqués, de détecter des erreurs de sens qui auraient pu se glisser à quelque moment de l'ensemble du processus de collecte.

Quant à la création de variables dérivées (indicateur du statut socioéconomique, perception du quartier, fonctionnement familial, tempérament du nourrisson, etc.), elle devient une autre occasion de vérifier la qualité des données soit par le biais de l'étude des distributions de fréquences (dans le cas de variables discrètes) ou des tracés graphiques (dans le cas de variables continues).

En dernier lieu, les chercheurs doivent signaler à la Direction Santé Québec toute donnée aberrante qu'ils repèrent lors de l'analyse. Ainsi des corrections finales pourraient être faites avant que la banque ne soit rendue disponible à d'autres utilisateurs.

Ceci clôt la présentation générale de l'ÉLDEQ 1998-2002 auprès des enfants de 5 mois. La prochaine section viendra compléter cette description en adoptant l'angle de la méthodologie statistique. Si plusieurs nouveaux aspects de cette étude sont abordés, certains points précédemment présentés feront l'objet d'une seconde évaluation mais, cette fois, d'un point de vue purement statistique. Les méthodologies d'enquête et statistique constituent, dans les faits, l'alpha et l'oméga de toute enquête d'envergure. Ces deux perspectives complémentaires deviennent indispensables à quiconque veut comprendre, utiliser, voire même reproduire, en tout ou en partie, les résultats de la présente enquête longitudinale.

3. Description de l'ÉLDEQ 1998-2002 auprès des enfants de 5 mois dans une perspective de méthodologie statistique

Comme mentionné précédemment, l'objectif premier de cette enquête est d'obtenir de l'information longitudinale. Cette information peut alors servir à effectuer des études sur l'évolution de plusieurs variables à travers le temps. Un second objectif de cette enquête est la production d'estimations transversales, c'est-à-dire de résultats représentatifs de la population à un temps donné. Cependant, ces résultats (ou estimations) sont uniquement produits pour la première année de l'enquête longitudinale étant donné que le devis d'enquête retenu ne tient pas compte du changement de la population dans le temps. En effet, pour effectuer des estimations transversales à chaque année, il aurait fallu ajouter annuellement, à l'échantillon longitudinal, de nouveaux arrivants tels que des enfants immigrants. Des contingences financières et le désir de créer une cohorte québécoise, non seulement comparable mais complémentaire à l'ELNEJ (l'enquête mère), ont amené la Direction Santé Québec et les chercheurs de l'ÉLDEQ 1998-2002 à opter pour un autre devis d'enquête.

Dans les prochaines sections, les principaux aspects méthodologiques de la première année de l'ÉLDEQ 1998-2002 sont décrits. De manière plus spécifique, la section 3.1 porte sur la méthodologie retenue pour la réalisation du projet. On y précise les populations visées et échantillonnées, la base de sondage, la stratification utilisée, la taille de l'échantillon ainsi que sa répartition. À la section 3.2, on traite de l'application de la méthodologie mise au point : le prétest de 1996, la collecte et les ajustements apportés aux données. L'évaluation statistique de l'enquête sera présentée à la section 3.3. Quant aux méthodes d'analyse transversale proposées, elles viennent clore ce numéro au point 3.4.

3.1 Méthodologie retenue

3.1.1 Les populations et la base de sondage

La population visée par l'enquête est l'ensemble des bébés (naissances simples seulement)¹¹ qui avaient 59 ou 60

semaines d'âge gestationnel¹² au début de chaque période de collecte, de mères vivant au Québec, à l'exclusion des mères vivant dans les régions sociosanitaires 10 (Nord-du-Québec), 17 (territoire cri) et 18 (territoire inuit) et des mères vivant sur des réserves indiennes. Selon le *Fichier maître des naissances 1997-1998*, ces exclusions représentent 2,1 % de toutes les naissances vivantes des mères résidant au Québec. Quant aux naissances pour lesquelles la durée de la grossesse est manquante dans le fichier (1,3 %), elles n'ont pas été conservées. Ainsi, la population visée à cette étape du processus représente environ 96,6 % de la population totale.

Le *Fichier maître des naissances* du MSSS contient les enregistrements de tous les formulaires de naissance par année civile. Ces formulaires indiquent par exemple le nom, la date de naissance, la langue maternelle du père et de la mère, l'adresse, la langue d'usage à la maison et le niveau de scolarité de la mère, la durée de la grossesse ainsi que certaines informations sur le bébé. La saisie et la mise à jour des formulaires s'effectuent quotidiennement. La fermeture du fichier a lieu approximativement au mois d'août de l'année suivant l'année civile visée.

Étant donné que la collecte de 1998 a eu lieu durant l'année civile en cours, l'échantillonnage a été effectué à partir d'une base de sondage en constante évolution. En effet, les données du *Fichier maître des naissances* sont sujettes à certaines variations : il peut y avoir des retards au moment de l'expédition des formulaires par les déclarants ou encore lors de leur codification. Ainsi, certains bébés qui auraient pu répondre aux critères de sélection pouvaient ne pas se trouver dans le fichier au moment du tirage. Selon le *Fichier maître complet des naissances 1997-1998*, cette sous-couverture est de l'ordre de 4,7 % de la population visée. Toutefois, il est important de mentionner que 50 % des bébés non enregistrés rovenaient de mères demeurant dans la région de Hull

11. Les couples de jumeaux (naissances gémellaires) et les autres naissances multiples ne sont pas visés par l'enquête.

12. L'âge gestationnel est défini comme étant la somme de la durée de gestation (de la grossesse) et l'âge chronologique du bébé. Du à la variation de la durée de grossesse et aux 4 à 5 semaines allouées à chaque vague de collecte, tous les nourrissons n'ont pas tout à fait le même âge au moment de la collecte.

(région sociosanitaire de l'Outaouais) qui avaient accouché en Ontario¹³. Une poststratification appropriée a permis de réduire un éventuel biais causé par cette sous-couverture¹⁴.

Certains bébés prématurés n'ont pu être retenus, soit ceux pour lesquels la durée de gestation s'élevait à moins de 24 semaines, à cause du risque plus élevé qu'un décès ne survienne entre l'enregistrement du bébé au fichier et la tenue de l'enquête. De la même façon, les naissances s'étant produites après 42 semaines de gestation durent être mises de côté en raison des délais pour le tirage qu'aurait entraîné l'attente de leur disponibilité dans la base de sondage. Ces deux derniers cas d'exclusion *a priori* représentent approximativement 0,1 % des naissances enregistrées à la date de collecte. Quant aux bébés pour lesquels la variable « sexe » était manquante (0,1 %), ils ont également été exclus. En dernier lieu, une légère sous-couverture (évaluée à 0,6 %) due à la variabilité des durées de grossesse admissibles à chaque vague de collecte est discutée à la section 3.1.3.

Ainsi, on estime donc que la population échantillonnée pour l'ÉLDEQ 1998-2002 représente environ 94,5 % de la population visée.

3.1.2 Le plan de sondage et la stratification

Le plan de sondage de l'enquête est stratifié à trois degrés. Le territoire visé par l'enquête est premièrement divisé en unités primaires d'échantillonnage (UPE) formées d'une région sociosanitaire. Ces régions sont ensuite partagées en deux groupes : régions « éloignées » et « non éloignées »¹⁵. Au premier degré, on échantillonne deux régions sociosanitaires « éloignées » parmi quatre; les 11 régions « non éloignées » sont automatiquement choisies. Cette dernière opération a comme objectif de maximiser le rendement des sommes investies dans la

collecte en restreignant le territoire à parcourir pour avoir une enquête populationnelle de qualité. Comme les quatre régions « éloignées » ne représentent que 6,8 % de la population visée, en conséquence l'impact de cette stratégie devrait être minime sur les résultats globaux de l'enquête.

Au deuxième degré, les régions sociosanitaires sélectionnées sont divisées en USE (unités secondaires d'échantillonnage) formées d'une ou deux municipalités régionales de comté (MRC). À l'intérieur de chaque région sociosanitaire, les USE sont divisées en deux groupes selon le nombre de naissances enregistrées en 1996¹⁶. Le premier contient les USE dont le nombre de naissances enregistrées est faible; parmi ces USE, on en sélectionne un nombre fixe avec une probabilité proportionnelle au nombre de naissances visées sur le territoire. Le second groupe se distingue par les USE dont le nombre de naissances enregistrées est élevé; toutes ces USE sont sélectionnées dans l'échantillon. Parmi les USE dont le nombre de naissances enregistrées est faible, on a choisi simultanément : 1) de fixer la sélection, pour les six vagues de collecte, des USE dans huit régions sociosanitaires et 2) de faire varier la sélection des USE, à chaque vague, pour quatre régions sociosanitaires (Saguenay–Lac-Saint-Jean, Estrie, Chaudière-Appalaches et Montérégie). Le fait de ne pas avoir fixé définitivement le choix de toutes les USE d'une vague à l'autre permet de contenir l'effet de plan global tout en affectant les coûts de collecte que de manière très marginale.

Au troisième et dernier degré, un nombre de bébés (UTE = unités tertiaires d'échantillonnage) est sélectionné à l'intérieur des USE.

Le tableau 3.1 présente la répartition de la population visée des nouveau-nés selon la région sociosanitaire et le sexe. La lecture de ce tableau permet également de prendre connaissance de la taille et de la répartition de l'échantillon dont il est question à la section suivante.

13. L'information concernant ces naissances est transmise au Québec par l'Ontario de façon sporadique et tardive.

14. Lors de cette poststratification, on fait l'hypothèse que les nouveau-nés exclus de l'enquête sont semblables à ceux couverts par l'enquête.

15. C'est uniquement pour les besoins de l'enquête que les quatre régions Bas-Saint-Laurent, Abitibi-Témiscamingue, Côte-Nord et Gaspésie-Îles-de-la-Madeleine ont été désignées « éloignées ». La distance qui les sépare des grands centres, leur étendue géographique et leur très faible densité de population sont les critères qui ont permis leur regroupement sous un même vocable.

16. Seul le *Fichier maître complet des naissances 1996* était disponible pour déterminer la taille de l'échantillon par région sociosanitaire.

Tableau 3.1

Répartition de la population visée et de l'échantillon des nouveau-nés selon la région sociosanitaire et le sexe, 1998

Région sociosanitaire	%	Sexe		Population visée ^a	Sexe		Échantillon	%
		M	F		M	F	n	
*Bas-Saint-Laurent	2,0	91	97	188	n. sél.	n. sél.	n. sél.	n. sél.
Saguenay – Lac-Saint-Jean	3,6	173	163	336	53	49	102	3,5
Québec	7,2	325	342	667	110	118	228	7,8
Mauricie	6,0	287	265	552	85	83	168	5,7
Estrie	4,0	190	174	364	59	55	114	3,9
Montréal	27,8	1 329	1 236	2 565	397	365	762	25,9
Outaouais	4,4	209	199	408	61	53	114	3,9
*Abitibi-Témiscamingue	2,3	105	109	214	64	62	126	4,3
*Côte-Nord	1,4	62	69	131	n. sél.	n. sél.	n. sél.	n. sél.
*Gaspésie – Îles-de-la-Madeleine	1,1	47	50	97	34	32	66	2,2
Chaudière-Appalaches	5,1	232	241	473	76	74	150	5,1
Laval	4,9	243	209	452	77	61	138	4,7
Lanaudière	5,7	282	243	525	94	74	168	5,7
Laurentides	6,4	298	295	593	92	88	180	6,1
Montérégie	18,1	848	828	1 676	316	308	624	21,2
Nouveau-nés	100,0	4 721	4 520	9 241	1 518	1 422	2 940	100,0

a. Identifie la population visée au 29 janvier 1999. À cette date, la fermeture du *Fichier maître des naissances 1998* n'était pas encore complétée; elle était prévue en août 1999.

* L'astérisque identifie les régions dites « éloignées ».

« n. sél. » identifie les 2 régions « éloignées » qui n'ont pas été sélectionnées dans l'échantillon.

Source : Institut de la statistique du Québec, *ÉLDEQ 1998-2002*.

3.1.3 La détermination de la taille et de la répartition de l'échantillon

La taille de l'échantillon de l'enquête a été fixée en tenant compte de deux impératifs liés à toute étude de cette envergure : 1) permettre l'obtention de statistiques longitudinales globales fiables sur une période de 5 ans et 2) respecter les limites budgétaires fixées. Initialement, un échantillon de 2 800 nouveau-nés a été déterminé¹⁷. Compte tenu des résultats du prétest (Santé Québec, Jetté, Desrosiers et Tremblay, 1997, p. 14), un taux de réponse de 72 % et un effet de plan de 1,3¹⁸ étaient anticipés. Cependant, à la vague 3 et aux vagues subséquentes, un suréchantillonnage pour la région sociosanitaire de la Montérégie était nécessaire pour

mesurer l'impact de la tempête de pluie verglaçante de janvier 1998 sur des variables importantes de la cohorte. Ce faisant, l'échantillon de départ a été augmenté à 2 940 nouveau-nés.

Cet échantillon est réparti annuellement, on l'a vu, en six vagues qui représentent autant de périodes de collecte disjointes. Au départ, on souhaitait répartir également l'échantillon entre les vagues afin que les intervieweurs aient une tâche équivalente à chaque vague. Cependant, la tempête de pluie verglaçante de janvier 1998, qui a paralysé la grande région de Montréal et une grande partie de la Montérégie pendant des périodes allant d'une à quatre semaines, a obligé une réduction de 50 % de la taille de l'échantillon de la première vague en Montérégie. Pour compenser cette diminution, les échantillons des vagues subséquentes en Montérégie ont été légèrement augmentés¹⁹.

17. Cet échantillon assurait une marge d'erreur de 1,5 % pour une proportion de l'ordre de 10 % associée à un niveau de confiance de 95 % pour des statistiques applicables à l'ensemble du Québec.

18. L'effet de plan est le quotient de la variance estimée en tenant compte du plan de sondage par la variance obtenue à partir d'un plan de sondage aléatoire simple de même taille (voir plus loin dans le texte).

19. Augmentation d'environ 10 % des ménages à chaque vague pour une récupération complète des 50 % retranchés à la première vague.

Il était important de vérifier lors de chacun des tirages de l'échantillon que tous les bébés sélectionnés étaient nés après le 1^{er} octobre 1997. Cette contrainte permettait de s'assurer que la totalité des enfants choisis feraient leur entrée dans le système scolaire au cours de la même année scolaire. Afin de répondre à cette exigence, on a dû faire en sorte que la durée de grossesse admissible des mères faisant partie de l'enquête diffère selon la vague de collecte :

- vague 1 : durée de grossesse admissible pouvait varier entre 37 et 42 semaines;
- vague 2 : durée de grossesse admissible pouvait varier entre 33 et 42 semaines;
- vague 3 : durée de grossesse admissible pouvait varier entre 29 et 42 semaines; et
- vagues 4 à 6 : durée de grossesse admissible pouvait varier entre 24 et 42 semaines.

Conséquemment, une légère sous-couverture, principalement à la vague 1, est à noter pour les bébés dont la durée de gestation se situait entre 24 et 36 semaines (soit les bébés prématurés). La sous-couverture est estimée à 0,6 % et une poststratification adéquate permet de corriger²⁰ ce genre d'irrégularité.

Le pourcentage de chevauchement de l'échantillon de l'ÉLDEQ 1998-2002 et de celui de l'ELNEJ a été analysé lors du prétest de 1996 (Des Groseilliers, 1997, p. 244). Le faible pourcentage de chevauchement incite à considérer les bébés couverts par les deux enquêtes comme des non-répondants.

Pour sélectionner les deux régions sociosanitaires dites « éloignées », un tirage aléatoire proportionnel à la taille a été utilisé : les deux régions retenues sont l'Abitibi-Témiscamingue et la Gaspésie, tel qu'identifié au tableau 3.1. Au deuxième degré, les USE ont été sélectionnées, on l'a vu, avec probabilité proportionnelle au nombre de naissances dans les strates caractérisées par un nombre plus réduit de naissances et avec certitude dans les autres strates. Au troisième degré, la sélection

des bébés (UTE) se fait selon un tirage proportionnel au nombre de naissances enregistrées (voir annexe 2). Il est à noter que deux fois plus de bébés ont été sélectionnés dans les deux régions « éloignées », incluses dans l'enquête afin d'augmenter la précision globale.

3.2 Application de la méthodologie

3.2.1 Le prétest de 1996²¹

Une enquête pilote a eu lieu entre septembre et décembre 1996. Ses deux principaux objectifs étaient d'évaluer la faisabilité d'une enquête auprès d'une cohorte de bébés québécois et d'estimer la prévalence de quelques phénomènes chez ces bébés. Elle a permis en outre de vérifier la qualité de la base de sondage choisie, la mise en œuvre du processus de tirage d'un échantillon, les préoccupations liées au processus de collecte, la qualité des questionnaires et les étapes de production de résultats. À la suite de l'enquête pilote, plusieurs approches ont été modifiées afin d'augmenter le taux de réponse pour l'enquête longitudinale : jumelage avec la RAMQ de fichiers administratifs pour augmenter l'exactitude des adresses; prolongation de l'âge d'admissibilité du bébé; diminution de la durée de l'entrevue; et ajout d'un incitatif annuel aux ménages participants.

3.2.2 Les taux de réponse de l'ÉLDEQ 1998-2002 auprès des nourrissons de 5 mois

Le tableau 3.2 présente la répartition des résultats de la collecte pour le QIRI, instrument qui, rappelons-le, doit nécessairement être rempli pour qu'un ménage soit considéré répondant. C'est après avoir rempli cet instrument que découle la participation des répondants aux autres instruments de l'enquête. À la lecture de ce tableau, on constate entre autres que le taux de réponse obtenu pour l'ÉLDEQ 1998-2002 auprès des enfants de 5 mois est de 75,8 %²².

Conformément à la définition de Hidiroglou, Drew et Gray (1993), le taux de réponse s'obtient par le rapport du

20. Apporter ce correctif suppose que les bébés prématurés exclus de l'enquête sont semblables à ceux qui y sont inclus.

21. Pour une description détaillée du prétest, consulter le rapport préliminaire publié par Santé Québec, Jetté, M. H. Desrosiers et R.E. Tremblay, 1997.

22. C'est-à-dire un taux de réponse de près de 4 % plus élevé que le taux anticipé (72 %) et de près de 15 % plus élevé que celui du prétest de 1996.

nombre d'unités répondantes (ménages participants) au nombre d'unités dans le champ de l'enquête (unités répondantes et non répondantes) et de cas non résolus.

Les unités répondantes sont les unités dans le champ de l'enquête qui ont répondu à la date limite de collecte de données. Les cas non résolus représentent les unités dont on ne peut établir l'état à la fin de l'étape de la collecte de données.

Tableau 3.2
Répartition des résultats de la collecte de l'ÉLDEQ 1998-2002, année 1998, selon la participation au QIRI, 1998

	n	%
Ménage répondant au QIRI :	2 223	75,6
Entrevue complétée	2 223	75,6
Entrevue semi-complétée	0	0,0
Unités répondantes (ménages participants)	2 223	75,6
Refus du ménage	440	15,0
Participation à l'ELNEJ	3	0,1
Fait partie du prétest longitudinal	1	0,0
Unités non répondantes	444	15,1
Non retracé	140	4,8
Langue étrangère	81	2,8
Hors province	25	0,8
Pas de réponse après 8 tentatives	5	0,2
Absence prolongée	6	0,2
Pas de téléphone	2	0,0
Commerce	6	0,2
Cas non résolus	265	9,0
Sous-total des unités non répondantes + cas non résolus	709	24,1
Décès du bébé	5	0,2
Jumeau	1	0,0
Handicap physique	2	0,0
Unités hors du champ de l'enquête	8	0,2
Total / Taux de réponse à l'enquête (2 223/(2 223+444+265))	2 940	75,8

Source : Institut de la statistique du Québec, *ÉLDEQ 1998-2002*.

Quant au tableau 3.3, il présente la liste des principaux instruments de collecte de données, le nombre de répondants à chacun des instruments entièrement remplis qui ont fait l'objet d'un traitement statistique de données et le taux de réponse correspondant.

Tableau 3.3

Liste des instruments de collecte, selon le nombre de répondants et le taux de réponse associé, 1998

	Nombre de répondants	Taux de réponse
QIRI	2 223	75,8 %
QPRI	2 223	75,8 %
Dos. médicaux	2 184	74,5 %
Jeu 1,2,3 mains ⁽¹⁾	1 851	63,1 %
QAAM	2 146	73,2 %
QAAP	1 855	66,2 %
QAAPABS ⁽²⁾	45	45,9 %
Journal de bébé ⁽³⁾	nd	nd
IMF	2 221	75,8 %

- 12,7 % des nourrissons n'ont pas complété le jeu totalement (voir section 2.2).
- Questionnaire autoadministré du père s'adressant aux pères absents du ménage ayant un contact avec l'enfant au moins une fois par mois.
- Au moment de la parution de ce numéro, il n'est pas encore exclu que l'on puisse publier dès l'année 2000 les résultats de cet instrument qui demande un traitement de données particulièrement ardu et sophistiqué. Peut-être constituera-t-il le treizième numéro de la présente collection (volume 1), sinon les données seront disponibles dans le volume 2, soit en 2002.

Source : Institut de la statistique du Québec, *ÉLDEQ 1998-2002*.

À l'exception du QAAPABS, s'adressant aux pères biologiques absents du ménage, pour lequel le taux de réponse est inférieur à 50 % (voir plus loin, section 3.4), les taux de réponse oscillent entre 63,1 % et 75,8 % selon l'instrument.

Que l'on parle des taux de non-réponse globale, c'est-à-dire celui de l'enquête ou d'un instrument donné, ou de taux de non-réponse partielle, c'est-à-dire la non-réponse à une ou des questions d'un instrument, il est primordial de mentionner qu'ils peuvent engendrer des biais plus ou moins importants dans les estimations.

Le tableau 3.4 présente les taux de réponse pour certains regroupements d'individus dont les caractéristiques sont liées aux variables mesurées dans l'enquête. Afin de limiter les biais potentiels dus à la non-réponse, des corrections sont faites en recourant à des données auxiliaires disponibles dans le *Fichier maître des naissances* pour l'ensemble de la population visée et notamment pour les non-répondants. Dans cette optique, on utilise deux méthodes standards pour réduire ces biais, soit l'ajustement pour la non-réponse et la poststratification. Ces thèmes sont abordés à la section suivante.

Tableau 3.4

Présentation des taux de réponse selon certaines variables démographiques et administratives^a, 1998

Catégories	Taux de réponse	
Âge de la mère	Moins de 19 ans	65,5 %
	20-24 ans	73,9 %
	25-29 ans	77,6 %
	30-34 ans	77,6 %
	35 ans et plus	69,4 %
Scolarité de la mère	Études primaires	43,3 %
	Études secondaires	65,8 %
	Cégep	74,3 %
	Études universitaires	83,6 %
Langue d'usage à la maison ^b	Français	79,3 %
	Anglais	68,5 %
	Autres	45,9 %
Région sociosanitaire	Saguenay – Lac-Saint-Jean	87,4 %
	Québec	84,0 %
	Mauricie	86,2 %
	Estrie	82,3 %
	Montréal	63,2 %
	Outaouais	78,0 %
	Abitibi-Témiscamingue	79,3 %
	Gaspésie – Îles-de-la-Madeleine	85,2 %
	Chaudière-Appalaches	63,4 %
	Laval	75,8 %
	Lanaudière	76,7 %
	Laurentides	79,8 %
Montérégie		

a. Variables apparaissant au *Fichier des naissances vivantes* et distinctes des variables dérivées qui sont utilisées dans les 12 ou 13 numéros de la présente collection (vol. 1).

b. Les catégories sont exclusives.

Source : Institut de la statistique du Québec, *ÉLDEQ 1998-2002*.

3.2.3 La pondération

Afin de procéder à l'inférence des données de l'échantillon à la population visée, chaque unité répondante doit se voir associer un poids. Ce poids est le nombre de bébés visés par l'enquête et « représentés » par le répondant. Les différentes étapes de la pondération dans l'enquête sont au nombre de trois. Il s'agit du calcul des poids préliminaires correspondants à l'inverse de la probabilité de sélection des bébés²³, de l'ajustement des poids pour tenir compte de la non-réponse et de la poststratification (Cox et Cohen, 1985).

23. Pour une présentation détaillée du calcul des poids préliminaires, voir l'annexe 3.

La non-réponse

L'objectif général de l'ajustement pour la non-réponse et la poststratification consiste à réduire un éventuel biais dans les estimations afin d'obtenir des estimations plus fiables. Cependant, pour mettre en œuvre ces techniques, il est nécessaire de disposer d'informations complémentaires autant sur les non-répondants que sur l'ensemble de la population. Dans le cadre de la présente enquête, cette condition est remplie puisque le *Fichier maître des naissances* contient un ensemble de variables administratives pour tous les bébés de la population visée.

Il faut rappeler que ces méthodes d'ajustement des poids ne garantissent pas de manière absolue la disparition de tout biais dans les estimations et présentent certaines limites. En effet, on cherche à redresser l'échantillon des répondants pour tenir compte de certaines caractéristiques des non-répondants. Cependant, il est toujours possible que les non-répondants présentent des spécificités pour lesquelles aucune information auxiliaire ne serait disponible, spécificités qui affecteraient la réponse, donc les estimations si ces derniers avaient pu faire l'objet de l'enquête. Ainsi, pour que l'ajustement effectué soit efficace, il est important que l'information auxiliaire dont on dispose soit liée aux variables mesurées dans l'enquête, sinon l'impact de l'ajustement sur la réduction d'un biais potentiel sera négligeable.

En résumé, l'ajustement pour la non-réponse consiste à redresser l'échantillon des répondants par une modification de la pondération afin de le rendre, dans la mesure du possible, semblable à l'échantillon tiré initialement.

Pour l'ajustement de la non-réponse, la méthode de création de cellules de pondération a été utilisée. La technique employée est la modélisation par segmentation. Cette approche est fondée sur l'algorithme *CHAID* (*Chi-Square Automatic Interaction Detection*)²⁴ détection automatique d'interactions du chi-carré mis au point par Kass (1980)²⁴.

24. L'algorithme *CHAID* utilise des tests du chi-carré pour diviser les prédicteurs catégoriques et retenir la division la plus significative à chaque étape. Ainsi, cette dernière permet de traiter un grand nombre de covariables et, en raison de la structure en arbre de décision, il est facile d'établir les interactions (Brien, 1998; Tambay et coll. 1998).

Le concept général de la méthode est le suivant : on considère que la population peut être divisée en cellules homogènes dans chacune desquelles tous les individus ont la même probabilité de répondre s'ils sont échantillonnés et que les probabilités de réponse sont différentes entre les cellules.

Donc, pour réduire le biais dû à la non-réponse de manière importante, on doit former des cellules d'individus sélectionnés de façon judicieuse selon les critères définis précédemment.

Parmi les variables administratives provenant du *Fichier maître des naissances*, on a utilisé l'âge de la mère, la scolarité de la mère et la langue d'usage à la maison. Ces variables semblaient les plus liées au phénomène étudié c'est-à-dire la connaissance de précurseurs de l'adaptation en milieu scolaire. La variable de stratification région sociosanitaire a également été utilisée (voir tableau 3.4). Le processus d'ajustement a conduit à la formation de 13 cellules sur l'ensemble de l'échantillon (voir l'annexe 4).

La poststratification

La poststratification consiste à stratifier la population après la collecte des données; elle requiert la connaissance de la proportion de la population appartenant à certains groupes (Cox et Cohen, 1985). Tout comme la stratification, la poststratification augmente la précision des estimations; elle contribue simultanément à diminuer un éventuel biais de non-réponse et à corriger pour la sous-couverture.

Tout comme pour la non-réponse, la poststratification entraîne un ajustement de la pondération. Elle consiste à ajuster la pondération afin de s'assurer que la distribution pondérée des répondants soit conforme à la distribution de la population visée ou, au besoin, au *Fichier maître des naissances 1998*, en fonction de certaines variables sociodémographiques. Les variables qui ont servi à l'ajustement sont : la région sociosanitaire, le sexe du bébé et la durée de la grossesse. Des groupes homogènes, appelés poststrates, sont formés à partir de ces variables (voir l'annexe 5).

3.3 Évaluation de l'enquête

Dans toute enquête statistique, les estimations produites sont entachées d'erreurs. D'une part, il y a les erreurs reliées à l'échantillonnage, c'est-à-dire celles dues au fait que l'on enquête seulement une partie de la population observée. D'autre part, il y a les erreurs dites « non dues à l'échantillonnage » pouvant être occasionnées par un taux de réponse jugé insuffisant, une faiblesse de la base de sondage, des questions difficiles à interpréter, des erreurs de saisie, etc.

3.3.1 Les erreurs dues à l'échantillonnage

On appelle erreur d'échantillonnage, la différence entre les estimations tirées d'un échantillon et les paramètres résultant d'un recensement de la population entrepris dans les mêmes conditions. Deux mesures de l'erreur d'échantillonnage sont utilisées dans l'enquête : la marge d'erreur et le coefficient de variation (CV).

La marge d'erreur, pour un risque de se tromper de 5 %, est une mesure de précision de l'estimation et se définit comme suit :

$$\text{marge d'erreur} = 1,96 \times \sqrt{\text{variance de l'estimation}}$$

À partir de la marge d'erreur, il est possible de définir l'**intervalle de confiance** (IC) à 95 % qui est associé à l'estimation :

$$\text{IC} = \text{estimation} \pm \text{marge d'erreur.}$$

Cet intervalle illustre l'étendue des valeurs possibles que peut prendre la variable étudiée dans la population observée. Cela signifie que, si l'on refaisait l'enquête à plusieurs reprises, 19 intervalles sur 20 contiendraient la valeur réelle de l'estimation.

Quant au **coefficient de variation (CV)**, il permet de quantifier la précision de l'estimation et se calcule de la façon suivante :

$$\text{CV} = \frac{\text{Marge d'erreur}}{(1,96 \times \text{estimation})}$$

Cette mesure contribue à faciliter l'interprétation de la précision d'une estimation. Plus le CV est élevé, moins précise est l'estimation et inversement. Le tableau 3.5

qui rend compte de la relation entre le CV et la précision de l'estimation, montre, entre autres, qu'il faut être prudent (apparition d'un astérisque * dans les cellules des tableaux du rapport) lors de l'utilisation d'estimations dont le CV est considéré « passable ». Quant aux estimations dont le CV est très élevé (apparition de deux astérisques ** dans les cellules des tableaux), elles sont sujettes à une très grande variabilité et ne doivent être utilisées qu'à titre indicatif. Le CV a également l'avantage de permettre de comparer la précision de différentes estimations entre elles.

Tableau 3.5
Qualité des estimations (utilisée par la Direction de la méthodologie et des enquêtes spéciales de l'ISQ), 1998

CV	Précision de l'estimation
≤ 5 %	Très bonne
5 % < CV ≤ 15 %	Bonne
15 % < CV ≤ 25 %	Passable *
> 25 %	Faible **

Source : Institut de la statistique du Québec, *ÉLDEQ 1998-2002*.

La précision de certaines proportions est présentée aux tableaux 3.6 et 3.7.

Tableau 3.6
Précision de certaines proportions (lorsque le nombre de répondants à une question atteint 2 223 ou 100 % des répondants de l'enquête), 1998

Proportion	Marge d'erreur	CV	Approximation utilisée
%			
1	0,6	24,1	Binomiale
2	0,8	16,9	Binomiale
2,5	0,9	15,0	Binomiale
3	0,9	13,8	Binomiale
4	1,0	11,9	Binomiale
5	1,0	10,5	Normale
10	1,4	7,3	Normale
15	1,7	5,8	Normale
20	1,9	4,8	Normale
30	2,2	3,7	Normale
50	2,4	2,4	Normale
70	2,2	1,6	Normale
80	1,9	1,2	Normale
95	1,0	0,6	Normale

Source : Institut de la statistique du Québec, *ÉLDEQ 1998-2002*.

Une mise en garde est de rigueur quant à l'estimation de l'intervalle de confiance (IC) des petites proportions

(Cochran, 1977). En effet, dans ce cas, l'approximation normale pour le calcul de l'IC n'est plus valide. On suggère l'approximation basée sur la distribution binomiale et calculée à l'aide de la distribution de F (Korn et Graubard, 1998).

Tableau 3.7
Précision de certaines proportions (lorsque le nombre de répondants à une question atteint 1 100 répondants de l'enquête, soit le nombre de répondants lors d'estimations par sexe), 1998

Proportion	Marge d'erreur	CV	Approximation utilisée
%			
2 %	1,2 %	25,0 %	Binomiale
5 %	1,7 %	15,0 %	Binomiale
10 %	2,0 %	10,3 %	Normale
15 %	2,4 %	8,2 %	Normale
20 %	2,7 %	6,9 %	Normale
30 %	3,1 %	5,3 %	Normale
50 %	3,4 %	3,4 %	Normale
70 %	3,1 %	2,3 %	Normale
80 %	2,7 %	1,7 %	Normale
95 %	1,7 %	0,8 %	Binomiale

Source : Institut de la statistique du Québec, *ÉLDEQ 1998-2002*.

Quant à l'efficacité du plan de sondage retenu pour l'enquête, elle peut être évaluée *a posteriori*. Pour ce faire, on compare les résultats obtenus (en matière de précision statistique) avec ceux qui auraient été obtenus si on avait appliqué un plan de sondage aléatoire simple. Cette comparaison s'effectue à l'aide de l'effet de plan. Pour l'enquête, l'effet de plan moyen pour des proportions est estimé à 1,3. Ainsi, l'estimation de la variance des proportions est environ 30 % plus grande que celle d'un plan aléatoire simple.

3.3.2 Les erreurs non dues à l'échantillonnage

Toutes les autres formes d'erreurs (erreurs d'observation, de réponse, de traitement et de non-réponse) sont des erreurs non dues à l'échantillonnage. Le type d'erreur qui peut avoir un impact important sur les résultats est l'erreur due à la non-réponse. En effet, la non-réponse peut induire des biais dans les résultats.

La non-réponse est totale (ou globale) lorsque, pour une raison ou une autre, une personne refuse de participer soit à l'enquête, soit à tout un instrument de l'enquête. À l'exception du QAAPABS, les taux de réponse globaux

oscillent entre 63,1 % et 75,8 %, selon l'instrument (voir tableau 3.3).

La non-réponse est partielle lorsque seulement une partie d'un questionnaire a été remplie. Il est habituel de considérer qu'un taux de non-réponse partielle inférieur à 5 % ne devrait pas susciter d'inquiétude. Toutefois, lorsque ce taux est supérieur à 5 %, des mises en garde doivent être faites à l'égard de la présence de biais potentiels. Une analyse de la non-réponse partielle des divers instruments a donc été effectuée. Elle consiste à évaluer la non-réponse pour chaque question, à cibler les cas problématiques et à déterminer l'importance de la non-réponse partielle à l'ensemble de chaque instrument. Les 6 questionnaires papier (QPRI, QAAM, QAAP, verglas, IMF, « Jeu 1,2,3 mains ») ainsi que le questionnaire informatique rempli par l'intervieweur (QIRI) ont été étudiés. À noter que le taux de non-réponse à chaque question est obtenu par le rapport entre le nombre pondéré d'individus n'ayant pas répondu à la question et le nombre pondéré d'individus devant répondre à cette question (répondants potentiels). En ce qui a trait aux taux de non-réponse partielle des « dossiers médicaux » et des « journaux de bébé », ils seront présentés dans le numéro qui leur sera respectivement consacré, le traitement de ces données complexes n'étant pas complété.

La plupart des questions de l'enquête affichent des taux de non-réponse partielle très bas, sauf pour certaines questions du QAAM, du QIRI et du QPRI (voir l'annexe 6 pour les mises en garde concernant les taux de non-réponse partielle excédant 5 %). Ainsi, dans la présente enquête, la non-réponse partielle ne semble pas être un problème important à l'égard de la présence de biais potentiels.

3.4 Méthodes d'analyse

La majorité des méthodes utilisées pour l'analyse de données ainsi que les logiciels permettant de les réaliser correspondent à un échantillonnage aléatoire simple. Comme on l'a vu précédemment, la présente enquête a un plan de sondage stratifié à trois degrés. Lors de l'analyse de données, si on fait l'hypothèse que l'échantillonnage est aléatoire simple, on prend le risque d'engendrer un biais dans les estimations et de sous-estimer leur variance. En conséquence, il peut en résulter des conclusions erronées au plan des tests statistiques.

Il est donc important de tenir compte du plan de sondage dans les analyses de données. Deux corrections complémentaires sont proposées : la première est liée aux estimations et la seconde concerne la précision des estimations.

La première correction à effectuer est la pondération des estimations. Comme on l'a vu à la section 3.2.3, la pondération consiste à attribuer à chaque répondant une valeur (un poids) correspondant au nombre de personnes qu'il « représente » dans la population. Elle permet ainsi l'inférence des données de l'échantillon à la population visée. Trois séries de poids transversaux ont été calculés : une pour le questionnaire autoadministré du père présent dans le ménage (QAAP), une autre pour le questionnaire papier rempli par l'intervieweur (QPRI) et une troisième pour le « Jeu 1, 2, 3 mains ». À noter que les poids obtenus pour le QPRI sont aussi valables pour les instruments QIRI, QAAM, IMF et les « Dossier médicaux » même si le nombre de répondants de certains instruments n'atteint pas 2 223. Pour les trois séries de poids, les poids populationnels et les poids échantillonnax sont fournis²⁵.

Notons cependant que le questionnaire autoadministré des pères absents ayant un contact avec l'enfant cible au moins une fois par mois (QAAPABS) n'a pas été pondéré, dû à son faible taux de réponse global (46 %), et au fait que les pères absents qui sont non-répondants semblent afficher un profil différent des pères absents qui sont répondants en regard du niveau de scolarité et du nombre de contacts avec l'enfant. De surcroît, ces derniers semblent différer des pères présents qui ont répondu au QAAP en regard de certaines caractéristiques mesurées dans le questionnaire. Ainsi, l'inférence des résultats à la population des bébés dont le père est absent est impossible. Cet échantillon permet toutefois d'effectuer certaines analyses descriptives qui auront trait uniquement aux bébés de l'échantillon dont le père est absent (voir le numéro 2 de la présente collection).

La seconde correction à réaliser touche la précision des estimations. Lorsqu'on ne possède pas les paramètres détaillés du plan de sondage, on peut déterminer la précision des estimations de proportion en utilisant une

25. Les poids échantillonnax sont des poids à moyenne 1 qui conservent le rapport entre les poids populationnels et dont la somme concorde avec le nombre de répondants à l'instrument.

estimation de l'effet de plan moyen ainsi que discuté à la section 3.3.1. On multiplie alors l'estimation de l'effet de plan moyen (1,3) par la variance d'un plan aléatoire simple pour obtenir la variance correspondant au plan de sondage de l'enquête. L'utilisation de l'effet de plan moyen est aussi appropriée pour éprouver la comparaison de deux proportions et pour les tests du chi-carré. Pour ces derniers, la correction choisie consiste à diviser la statistique par l'effet de plan moyen. Cette opération est équivalente à diviser les poids à moyenne 1 par l'effet de plan moyen.

Cependant pour les cas litigieux, c'est-à-dire lorsque le seuil observé du test est proche du seuil fixé pour le test du chi-carré ou lorsque les intervalles de confiance servant à la comparaison de proportions se chevauchent à peine, il faut employer une autre approche que celle de l'effet de plan moyen. Pour ces cas, il est suggéré d'utiliser un logiciel qui permet de calculer des estimations de variance en tenant compte d'un plan de sondage complexe tel que SUDAAN²⁶ (SAS, 1985; Shah, Barnwell, Hunt et Lavange, 1992).

Pour les analyses multivariées, aucune méthode approximative n'est proposée. Il est suggéré, dans un premier temps, de procéder sans tenir compte du plan de sondage au moment des estimations de variances, mais tout en pondérant avec un poids à moyenne 1 (poids échantillonaux). Dans un deuxième temps, les modèles finaux devraient être repris avec SUDAAN. Aucune variable ne serait ajoutée au modèle; cependant certains paramètres pourraient ne plus être significatifs. Parmi les analyses multivariées, SUDAAN permet d'effectuer des régressions linéaires et logistiques, des modèles log-linéaires et des analyses de variances.

3.4.1 Portée et limites des données

Après avoir présenté en détail les fondements méthodologiques de l'ÉLDEQ 1998-2002, que ce soit d'un point de vue statistique ou non, le lecteur est en mesure de faire une interprétation et une utilisation judicieuse des données. En rappel, on mentionne :

- 1) l'enquête comporte huit principaux instruments de collecte;
- 2) compte tenu de la variation des taux de réponse aux différents instruments, trois séries de poids doivent être calculés et doivent être utilisés pour que l'inférence à la population soit fiable;
- 3) à part le QAAPABS et une série de questions du QIRI concernant les pères biologiques absents – deux instruments dont la non-réponse globale ou partielle est trop élevée – tous les instruments ont pu être pondérés et, de fait, les estimations présentées ont toutes fait l'objet d'ajustements visant à réduire les biais potentiels;
- 4) toutes les données qui seront affectées par un coefficient de variation (CV) de plus de 15 % seront accompagnées d'un ou deux astérisques pour bien indiquer aux lecteurs la variabilité des estimations présentées;
- 5) si des estimations affectées par un taux de non-réponse partielle dépassant les 5 % sont présentées, elles seront accompagnées d'une note qui spécifie pour quel sous-groupe de la population elles sont moins fiables.

Si comme toute mesure transversale d'une enquête populationnelle, le volet *Nourrissons de 5 mois* de l'ÉLDEQ 1998-2002 présente certaines limites, il faut tout de même retenir, qu'à titre de première mondiale, cette mesure transversale demeure d'excellente qualité. La très grande majorité des estimations qu'elle procure sont valides et fiables et fournissent pour la première fois un portrait particulièrement détaillé des nourrissons de 5 mois, nés au Québec en 1997-1998. Loin d'être leur seul mérite, les enquêtes innovatrices doivent être réalisées une première fois pour ensuite faire l'objet d'améliorations. Cette mesure initiale de la cohorte annuelle québécoise a dépassé les espérances de tous nos partenaires : taux de réponse en hausse de plus de 15 % comparativement au prétest; taux de non-réponse partielle très majoritairement en deçà des 5 %; pratiquement tous les instruments ont été remplis totalement, donc aucune nécessité de recourir à l'imputation; des répondants reconnus comme difficiles dans la littérature – les hommes en général – et de nouveaux participants pratiquement inconnus du monde des enquêtes populationnelles et les

26. Pour le test du chi-carré, la correction de Satterthwaite dans SUDAAN devrait être utilisée. Pour les tests d'égalité de proportion, SUDAAN mesure la précision associée à la différence des proportions à l'aide de la méthode de linéarisation de Taylor.

nourrissons de 5 mois ont tous montré un haut taux de participation. En résumé, la qualité des estimations présentées dans les 12 (possiblement 13) numéros de cette collection initiale est établie et c'est avec enthousiasme que l'on envisage le suivi longitudinal de cette première cohorte annuelle québécoise.

