



Étude de l'attrition dans l'ÉLDEQ de 1998 à 2008

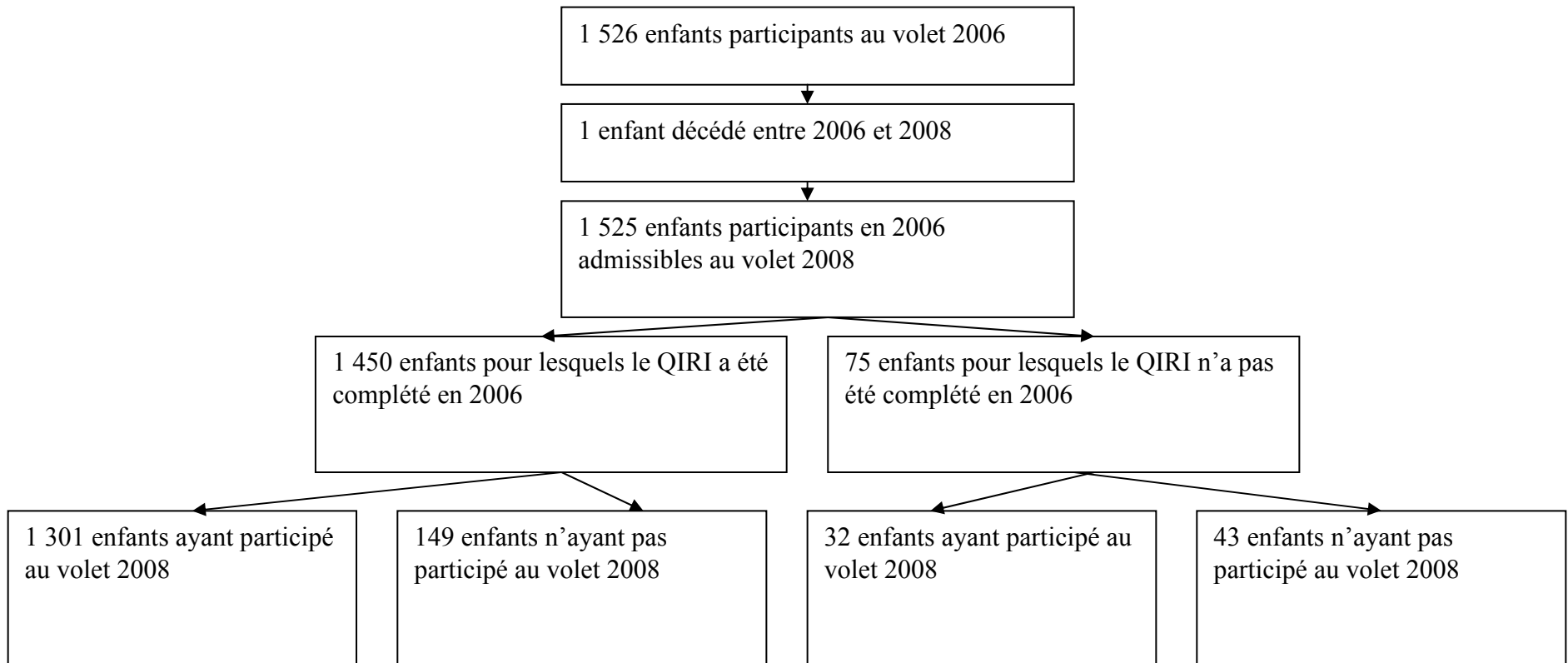
Affiche présentée dans le cadre du
Symposium de l'ÉLDEQ

par

Catherine Fontaine et Robert Courtemanche
Direction de la méthodologie et de la qualité
Institut de la statistique du Québec

11 novembre 2008

Participation à l'ÉLDEQ en 2008 versus 2006



Note: 60 enfants ont participé au volet 2008 mais n'avaient pas participé au volet 2006. Ils ne figurent pas dans cet organigramme.



Qui sont les non-participants au volet 2008 de l'ÉLDEQ?

On étudie la participation au volet 2008 parmi les familles ayant complété le QIRI en 2006 (n=1450).
La non-réponse partielle est de l'ordre de 1% ou moins pour les variables étudiées.

Caractéristiques des enfants participants en 2006	Taux de non-participation au volet 2008 (non pondéré)
Les revenus du ménage sont insuffisants ou très insuffisants (n=218)	17% (versus 9% pour les autres enfants)
La langue parlée par la mère est autre que le français seulement (n=230)	17% (versus 9% pour les autres enfants)
La mère est immigrante (n=117)	16% (versus 10% pour les autres enfants)
La principale source de revenus du ménage est autre qu'un salaire ou traitement (n=231)	16% (versus 9% pour les autres enfants)
La mère a comme plus haut niveau de scolarité un D.E.S. ou moins (n=266)	14% (versus 9% pour les autres enfants)
La mère ne travaille pas au moment de l'enquête (n=334)	13% (versus 9% pour les autres enfants)

Note : La participation au volet 2008 est définie comme le fait d'avoir complété le Questionnaire informatisé rempli par l'intervieweur (QIRI) ou le Questionnaire papier administré à l'enfant (QPAE).

Concept d'attrition et méthodes utilisées

A) Définition de l'attrition

Les analyses qui suivent portent sur l'attrition, définie comme la non-réponse à un volet donné de 1999 à 2006 (temps 2 à 9) parmi les répondants au volet 1998 (temps 1).

Cette définition a été retenue pour refléter le fait que les analyses effectuées portent le plus souvent sur un ou quelques volets d'enquête à la fois. Cette définition permet donc de réaliser une étude qui concorde le plus possible avec l'utilisation qui est faite des données longitudinales.

B) Méthodes utilisées

Les méthodes utilisées pour l'étude d'attrition sont:

- 1) Graphiques permettant d'illustrer visuellement la cohérence longitudinale (deux graphiques sont présentés)
- 2) Différences relatives permettant de mesurer la cohérence longitudinale (non présentées sur cette affiche)

$$\text{Différence relative} = \frac{\left(\hat{p}_{\text{transversa le pondérée volet } x} - \hat{p}_{\text{transversa le pondérée volet 1998}} \right)}{\hat{p}_{\text{transversa le pondérée volet 1998}}}$$

pour x=1999 à 2006

- 3) Tableau de l'évolution du nombre de répondants aux volets 2006 et 2008 pour une caractéristique mesurée en 1998 (présenté sur cette affiche)



Étude de la cohérence longitudinale de 1999 à 2006

Le premier aspect de l'étude d'attrition porte sur la cohérence longitudinale. On définit une estimation comme étant cohérente au niveau longitudinal si sa valeur est près de la valeur de l'estimation pondérée obtenue au début de l'enquête (année 1998).

Pour différentes caractéristiques mesurées en 1998, la cohérence longitudinale a été évaluée pour les estimations transversales pondérées des volets 1999 à 2006.

Deux graphiques sont présentés: un où le niveau de cohérence longitudinal est très bon (niveau de suffisance de revenus) et un où il est moins bon (langue parlée par la mère).

La courbe en **bleu** représente l'évolution de l'estimation (pour une caractéristique donnée) pondérée à l'aide du poids initial (inverse de la probabilité de sélection). Cette courbe se rend jusqu'au volet 2008.

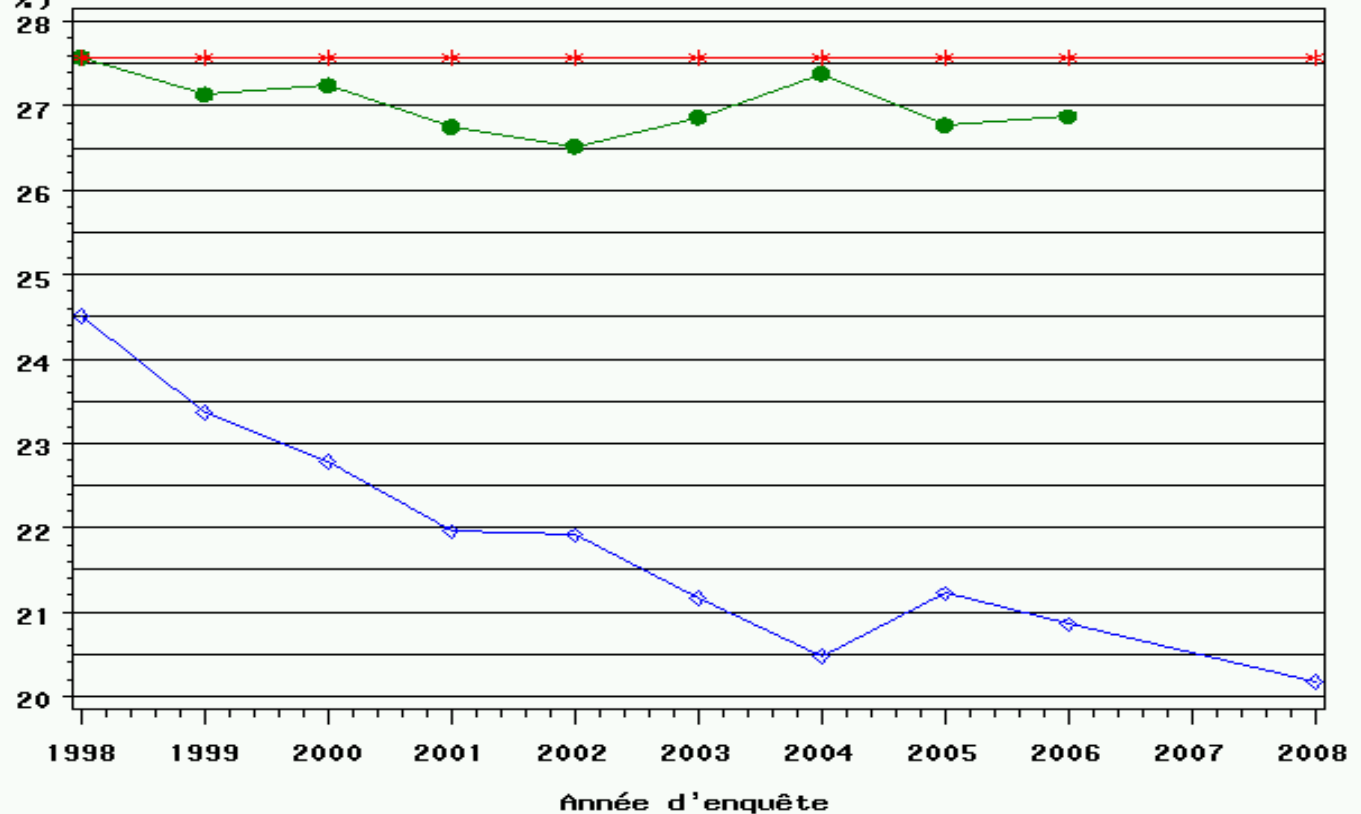
La courbe en **vert** représente l'évolution de l'estimation (pour une caractéristique donnée) pondérée pour tenir compte de la non-réponse survenue à un volet donné au niveau transversal. Cette courbe se rend jusqu'au volet 2006 (les données pondérées pour l'année 2008 ne sont pas disponibles).

Cette courbe peut être comparée à la droite en **rouge** qui représente l'estimation pondérée obtenue au début de l'enquête (année 1998). Cette estimation constitue la référence lorsque l'on veut étudier la cohérence longitudinale

Exemple: très bon niveau de cohérence longitudinale

Proportion estimée d'enfants provenant
de ménages ayant des revenus
insuffisants ou très insuffisants en 1998

Proportion d'enfants (en %)



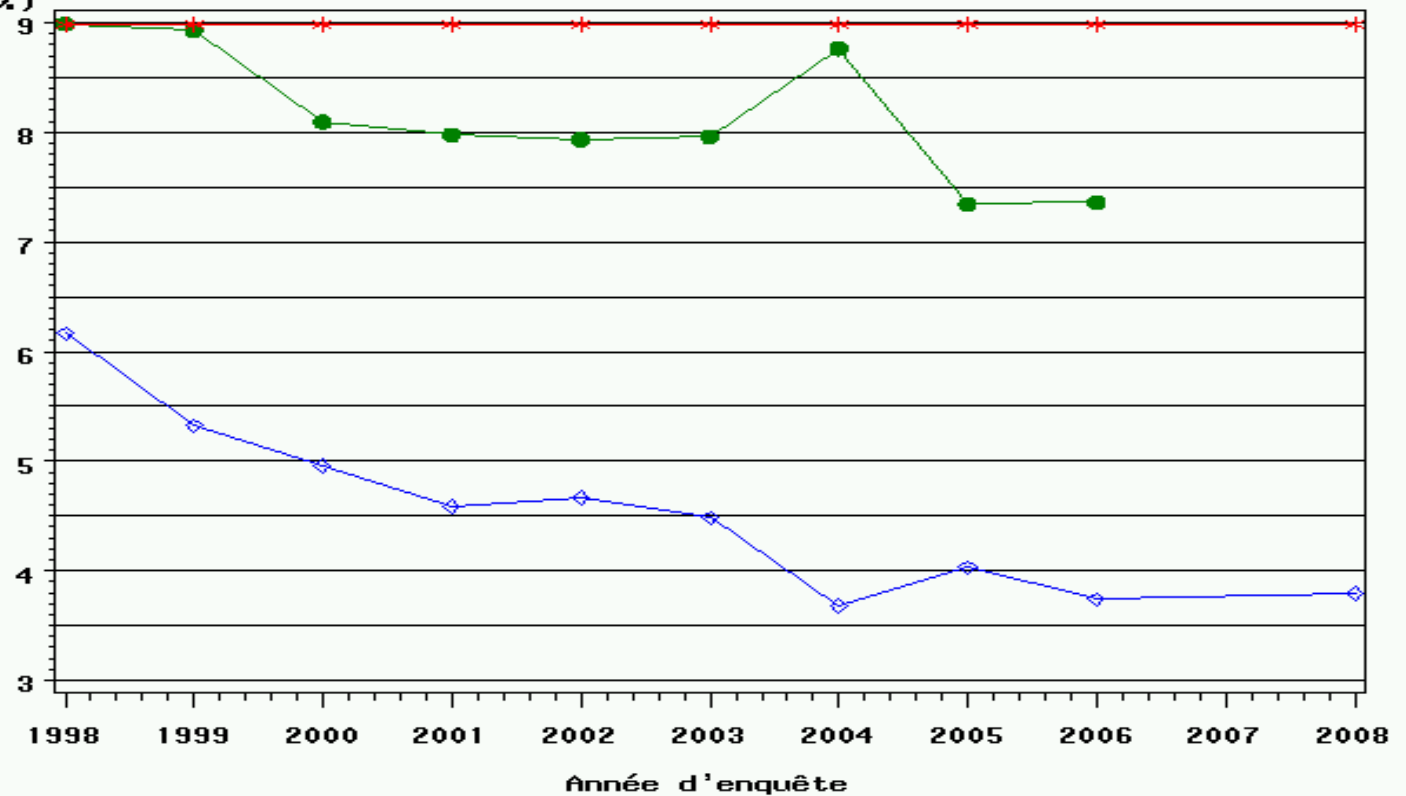
Courbes

- Estimations transversales (avec correction pour la non-réponse)
- *-*-* Estimation transversale de référence (année 1998)
- ◇-◇-◇ Estimations initiales (sans correction pour la non-réponse)

Exemple: moins bon niveau de cohérence longitudinale

Proportion estimée d'enfants
dont la mère ne parle ni français,
ni anglais en 1998

Proportion d'enfants (en %)



Courbes

- Estimations transversales (avec correction pour la non-réponse)
- *-*-* Estimation transversale de référence (année 1998)
- ◇◇◇ Estimations initiales (sans correction pour la non-réponse)



Résultat de l'analyse des différences relatives

NIVEAU DE COHÉRENCE LONGITUDINALE		
TRÈS BON	BON	MOINS BON
<p>>Enfants dont les revenus du ménage sont insuffisants ou très insuffisants en 1998</p> <p>>Enfants dont la santé en 1998 est bonne, passable ou mauvaise</p> <p>>Enfants de mères qui avaient moins de 25 ans en 1998</p>	<p>>Enfants dont les revenus du ménage sont faibles en 1998 (<15 000 \$)</p> <p>>Enfants de familles monoparentales en 1998</p> <p>>Enfants de mères ayant comme plus haut niveau de scolarité un D.E.S. (ou moins) en 1998</p> <p>>Enfants dont la région socio-sanitaire de résidence lors de l'échantillonnage est Montréal ou Laval</p>	<p>>Enfants de mères immigrantes en 1998</p> <p>>Enfants de mères qui ne parlent ni le français, ni l'anglais en 1998</p>

Constats:

- 1) Les 2 variables où le niveau de cohérence longitudinale est le moins bon, parmi les variables étudiées, sont des variables qui font habituellement l'objet d'un traitement de non-réponse. Par contre, il semble que même ces corrections ne permettent pas de rétablir un bon niveau de cohérence longitudinale.
- 2) Pour la plupart des variables étudiées, il semble plus difficile d'atteindre un bon niveau de cohérence longitudinale lors des volets récents (2004 à 2006).



Évolution du nombre de répondants: 1998 à 2008

Le deuxième aspect de l'étude d'attrition concerne l'évolution du nombre de répondants entre le volet 1998 et les volets 2006 et 2008. Les mêmes caractéristiques que celles évaluées pour l'étude de la cohérence longitudinale sont utilisées. En effet, l'attrition peut avoir un effet sur la cohérence longitudinale, de même que sur la précision des estimations et la puissance des tests.

Constats:

- 1) Un groupe de quatre variables a subi une perte variant entre 50% et 60% entre 1998 et 2008. Deux d'entre elles avaient un niveau moins bon de cohérence longitudinale (enfants de mères immigrantes et enfants de mères moins scolarisées en 1998).
- 2) Deux des trois variables qui avaient un très bon niveau de cohérence longitudinale présentent des pertes de répondants les plus faibles (environ 40%). Ce sont les variables représentant les enfants dont la mère a moins de 25 ans et les enfants dont la santé du bébé est bonne, passable ou mauvaise en 1998.
- 3) Les variables portant sur les enfants faisant partie de familles monoparentales et sur les enfants dont la mère ne parle ni français, ni anglais en 1998 ont un faible nombre de répondants en 2008 (moins de 90).

Évolution du nombre de répondants: 1998 à 2008 (fin)

Caractéristiques des enfants participants en 1998	Volet 1998	Volet 2006	Volet 2008	Perte entre 1998 et 2008 (non pondérée)
La mère ne parle ni le français, ni l'anglais	128	55	51	60%
La mère est immigrante	253	128	119	53%
Les revenus du ménage <15 000\$	248	140	119	52%
La famille est monoparentale	171	100	85	50%
Les revenus du ménage sont insuffisants ou très insuffisants	511	316	280	45%
La mère a comme plus haut niveau de scolarité un D.E.S. (ou moins)	564	370	327	42%
La région socio-sanitaire de résidence est Montréal ou Laval (au moment de l'échantillonnage)	571	371	334	42%
La santé de l'enfant est bonne, passable ou mauvaise	159	109	96	40%
La mère a moins de 25 ans	473	330	290	39%
Participation totale pour le volet	2 120	1 526	1 393	34%



Sans surprise, ce sont des caractéristiques liées au:

- a) Statut d'immigrant de la mère en 1998
- b) Plus haut niveau de scolarité de la mère en 1998
- c) Revenus du ménage en 1998
- d) Langue parlée par la mère à la maison en 1998
- e) Type de famille en 1998

qui subissent davantage les contrecoups de l'attrition dans l'ÉLDEQ. Si elles sont analysées et qu'elles sont également liées à la réussite scolaire, il y a un risque accru de biais longitudinal. Et ce, en dépit du fait que ce sont des caractéristiques qui sont habituellement prises en compte lors de l'ajustement pour la non-réponse effectué pour les différents instruments/volets. Enfin, elles sont généralement associées à de faibles nombres de répondants en 2008, ce qui peut entraîner des restrictions au niveau de l'analyse.

Que peut-on faire pour améliorer la situation lors de la suite de l'ÉLDEQ ?

- 1) Des travaux méthodologiques sont en cours afin d'évaluer une méthode de correction pour la non-réponse alternative à celle utilisée actuellement. Son effet sur le niveau de cohérence longitudinale est à déterminer. À suivre...
- 2) Si on veut étudier un phénomène portant sur **l'ensemble des enfants du Québec nés en 1998 (sauf les naissances multiples, les enfants nés hors-Québec et les enfants nés à des mois non couverts lors de l'échantillonnage)**, on peut vouloir ajouter un échantillon supplémentaire d'enfants à celui de l'ÉLDEQ. Les enfants seraient sélectionnés de manière à ce qu'ils possèdent des caractéristiques liées à l'attrition->grosso modo, des enfants provenant d'un milieu défavorisé ou des enfants ayant une mère immigrante. Voici des solutions à évaluer:

Solutions proposées	Quelques limites
Jumeler l'échantillon de l'étude « Québec en forme » à celui de l'ÉLDEQ. L'échantillon est constitué d'enfants provenant d'écoles situées en milieu défavorisé.	<ul style="list-style-type: none"> >Évaluer la faisabilité quant à l'appariement des deux échantillons (qualité, inférence, mesures) >L'étude « Québec en forme » contient des mesures rétrospectives sur l'enfant : sont-elles compatibles avec les mesures de l'ÉLDEQ?
Jumeler un échantillon qui serait sélectionné à partir du fichier administratif du ministère de l'Éducation, du Loisir et du Sport (MELS) du Québec à celui de l'ÉLDEQ. Les enfants présentant des caractéristiques reliées à l'attrition (écoles situées en milieu défavorisé; absence du père dans la famille au moment de l'inscription; langue d'usage à la maison autre que le français) sont visés par cet échantillon additionnel.	<ul style="list-style-type: none"> >Pas de données antérieures sur l'enfant à partir de l'ajout d'échantillon.



Idées pour la phase 3 de l'ÉLDEQ

Par ailleurs, si au cours de la phase 3 de l'ÉLDEQ il est prévu d'étudier un phénomène portant sur **l'ensemble des enfants du Québec qui sont nés en 1998**, il est à noter que certaines sous-populations ne sont pas couvertes présentement dans l'ÉLDEQ:

- a) Jumeaux et autres naissances multiples (environ 2% en 1998)
- b) Enfants nés hors du Québec (environ 9% en 2006)
- c) Enfants nés à un mois non couvert lors de l'échantillonnage en 1998 (environ 50% en 1998)

on peut également vouloir ajouter un échantillon supplémentaire d'enfants pour une (ou plusieurs) des sous-populations qui ne sont pas couvertes. Voici des propositions à évaluer:

Solutions proposées	Quelques limites
Jumeler l'échantillon de l'Étude des jumeaux nouveau-nés en 1998 à celui de l'ÉLDEQ.	<ul style="list-style-type: none"> >Évaluer la faisabilité quant à l'appariement des deux échantillons (qualité, inférence, mesures) >Beaucoup d'efforts pour évaluer la faisabilité par rapport à l'apport sur l'inférence (représente 2% de la population visée)
Jumeler un échantillon qui serait sélectionné à partir d'un fichier administratif à déterminer (Fichier maître des naissances vivantes de 1998, fichier provenant de la Régie de l'Assurance-Maladie du Québec ou fichier provenant du MELS) à celui de l'ÉLDEQ. Cette solution suppose que le fichier utilisé permette d'identifier une ou des sous-populations non couvertes par l'enquête initiale.	<ul style="list-style-type: none"> >Pas de données antérieures sur l'enfant à partir de l'ajout d'échantillon.



Mot de la fin

Merci de votre attention et bonne journée!

Catherine Fontaine
Statisticienne-méthodologiste
catherine.fontaine@stat.gouv.qc.ca

Robert Courtemanche
Coordonnateur en méthodologie
robert.courtemanche@stat.gouv.qc.ca