

CONFÉRENCE

Le croisement de données à des fins de statistiques et de recherches scientifiques

M^{me} Anne Monique Nuyt, MD, Service de néonatalogie, Centre de recherche, CHU Sainte-Justine

M^{me} Isabelle Leroux, analyste-conseil, Direction des services informationnels et technologiques, Institut de la statistique du Québec

Ces dernières années, un enjeu majeur en matière de résultats de recherche reçoit une attention accrue : l'accès, à des fins de recherche ou de gouverne, à des fichiers de microdonnées issus du couplage des fichiers administratifs ou d'enquêtes. De tels fichiers recèlent un potentiel analytique élevé. Ils permettent ainsi d'avoir des données probantes sur la santé d'une population grâce aux études populationnelles, de colliger de l'information sur un grand nombre de variables et de faciliter le suivi des cohortes dans le temps. Soucieux de contribuer au développement des connaissances pour la société québécoise, l'Institut de la statistique du Québec offre un ensemble de services qui permettent d'exploiter l'information existante dans les banques administratives détenues par les organismes gouvernementaux et d'apparier ces données entre elles ainsi qu'avec celles des banques de données des chercheurs.

M^{me} Isabelle Leroux a décrit brièvement comment l'Institut facilite l'accès à des microdonnées couplées, avec ou sans identifiant unique, tout en offrant les garanties requises de protection de la vie privée, conformément aux lois québécoises sur les renseignements personnels.

La D^{re} Anne Monique Nuyt a poursuivi en parlant de l'importance du croisement des données ainsi que des impacts présents et à venir sur la recherche en santé des populations. À cet effet, elle a présenté un de ses projets portant sur les risques de complications durant la grossesse chez les femmes nées prématurément. Cette étude a été réalisée grâce au couplage de données provenant du fichier des naissances du registre des événements démographiques et de Med-Écho, à partir de la plateforme de recherche de l'Institut de la statistique du Québec.

8 pages

Aux fins d'utilisation de cette documentation accessible aux membres via le site Internet de l'AAPI, nous vous prions de vous référer à l'énoncé des conditions d'utilisation des documents déposés dans la Médiathèque de la section réservée aux membres. Plus précisément, il est strictement interdit de modifier le contenu, par addition, suppression ou autrement, et d'en soustraire le nom ou le logo de l'AAPI ou de ses partenaires. Tous les documents mis à la disposition des membres appartiennent à l'AAPI ou à ses concédants de licence. Cette documentation vous est offerte à des fins d'information personnelle. L'AAPI et ses concédants de licence ne font aucune déclaration et ne donnent aucune garantie que le contenu est exhaustif, fiable, exact, exempt d'erreurs, de défauts, exempt de virus informatiques ou d'autres éléments nuisibles, ou que tous problèmes de cet ordre qui pourraient survenir sont corrigés. L'AAPI et ses concédants de licence ne sauraient être tenus responsables de quelque dommage que ce soit, y compris des dommages directs, indirects, accessoires, spéciaux, consécutifs, exemplaires ou punitifs découlant de l'accès à la documentation et aux services offerts à ses membres, de leur contenu ou de leurs outils, ou encore de l'accès à ceux-ci ou de l'utilisation de ceux-ci.

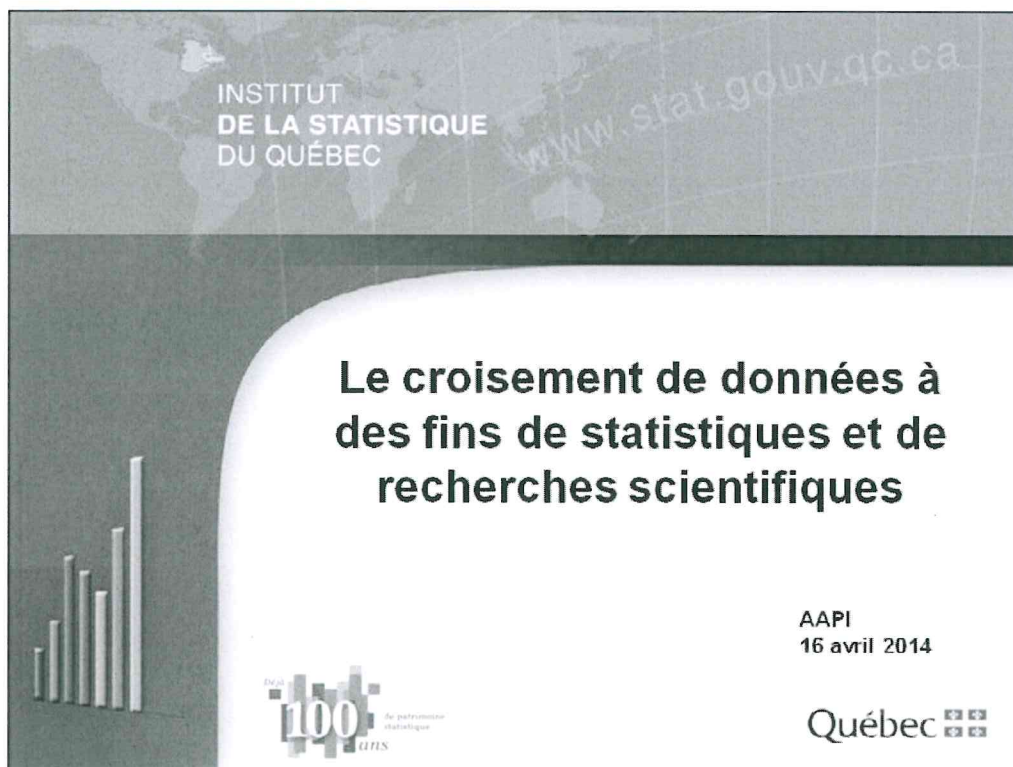
Vous pouvez reproduire ce texte, en tout ou en partie, en citant toutefois la source suivante : Nuyt, A. M. et Leroux, I. (2014). Conférence. « Le croisement de données à des fins de statistiques et de recherches scientifiques », Actes du 22^e Congrès AAPI 2014, Montréal, SOQUIJ, p. 60-67. Médiathèque AAPI - ©AAPI-CONGRÈS 2018-2019

CONFÉRENCE 4


Le croisement de données à des fins de statistiques et de recherches scientifiques

D^{RE} ANNE MONIQUE NUYT
Médecin, service de néonatalogie, CHU Sainte-Justine

M^{ME} ISABELLE LEROUX
Statisticienne, Institut de la statistique du Québec



www.stat.gouv.qc.ca




Plan de la présentation

- Rôle de l'Institut de la statistique du Québec (Institut)
- Les services d'accès aux données à des fins de recherche (SAD)
- Le processus d'accès aux données croisées
- Un exemple de projet réalisé nécessitant le croisement de données
- L'appariement (croisement) de banque de données

Institut de la statistique Québec

2

www.stat.gouv.qc.ca



Introduction

Rôles de l'Institut :

- ❖ Fournir des informations statistiques fiables et objectives sur la situation du Québec;
- ❖ Responsable de la réalisation de toutes les enquêtes statistiques d'intérêt général.

Rôle des services d'accès aux données à des fins de recherche (SAD) :

- ❖ Fournir un ensemble de services permettant de faciliter l'accès aux données statistiques d'enquêtes ou administratives à des fins de recherche scientifique, tout en offrant des garanties de confidentialité et de protection des renseignements personnels.

Institut de la statistique Québec

3

www.stat.gouv.qc.ca

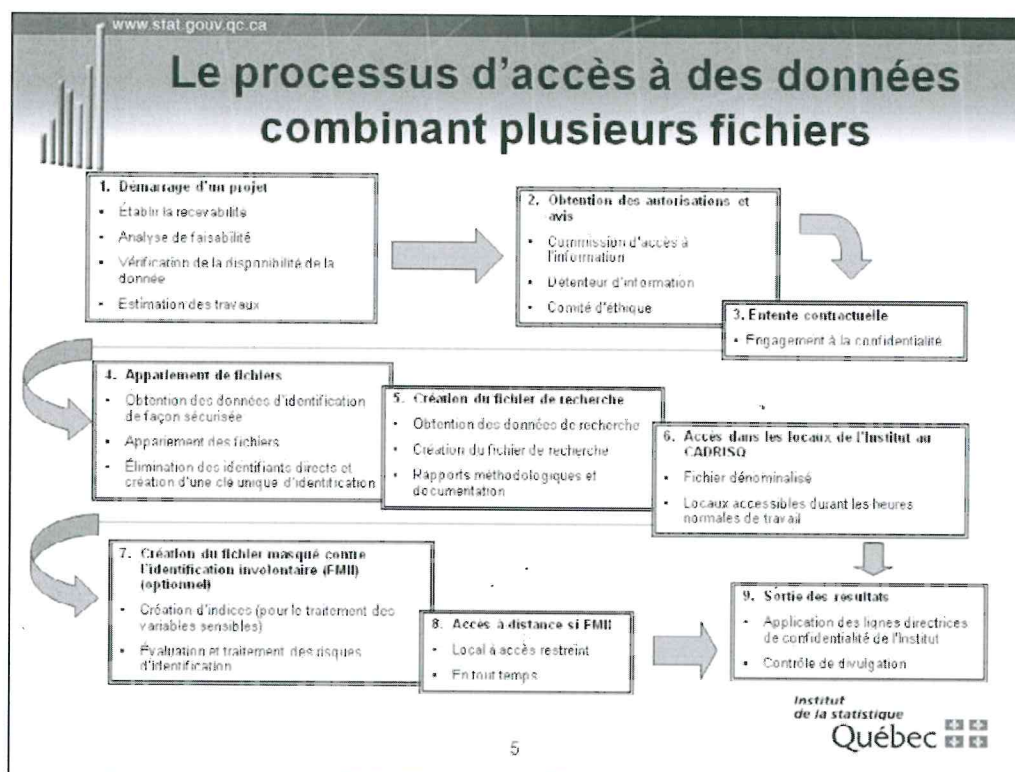
Les services d'accès aux données à des fins de recherche

Pour répondre aux besoins de la recherche, les SAD offrent aux chercheurs, étudiants et analystes de la fonction publique :


- ❖ Un accompagnement dans la démarche d'accès à des données d'enquêtes ou à des renseignements issus du Registre des événements démographiques;
- ❖ Une expertise en appariement de banques de données;
- ❖ Un accès aux données de recherche selon divers modes et la documentation nécessaire à la compréhension de ces données.

Institut de la statistique Québec

4



www.stat.gouv.qc.ca



Exemple de projet de recherche réalisé grâce au croisement de données

Institut de la statistique Québec

6

CMAJ RESEARCH

Pregnancy complications among women born preterm

Ariane Boivin PhD, Zhong-Cheng Luo MD PhD, François Audibert MD, Benoit Masse PhD, Francine Lefebvre MD, Réjean Tessier PhD, Anne Monique Nuyt MD

ABSTRACT

Background: Adults who were born with low birth weights are at increased risk of cardiovascular and metabolic conditions, including pregnancy complications. Low birth weight can result from intrauterine growth restriction, preterm birth or both. We examined the relation between preterm birth and pregnancy complications later in life.

Methods: We conducted a population-based cohort study in the province of Quebec involving 7405 women born preterm (554 < 32 weeks, 6851 at 32–36 weeks) and a matched cohort of 16 714 born at term between 1976 and 1995 who had a live birth or stillbirth between 1987 and 2008. The primary outcome measures were pregnancy complications (gestational diabetes, gestational hypertension, and preeclampsia or eclampsia).

Results: Overall, 19.9% of women born less than 32 weeks, 13.2% born at 32–36

weeks and 11.7% born at term had at least 1 pregnancy complication at least once during the study period ($p < 0.001$). Women born small for gestational age (both term and preterm) had increased odds of having at least 1 pregnancy complication compared with women born at term and appropriate for gestational age. After adjustment for various factors, including birth weight for gestational age, the odds of pregnancy complications associated with preterm birth was elevated by 1.95-fold (95% confidence interval [CI] 1.54–2.47) among women born before 32 weeks' gestation and 1.14-fold (95% CI 1.03–1.25) among those born at 32–36 weeks' gestation relative to women born at term.

Interpretation: Being born preterm, in addition to, and independent of, being small for

Competing interests: Benoit Masse has received consultancy fees from the International Partnership for Microbicides. No other competing interests were declared.

This article has been peer reviewed.

Correspondence to: Anne Monique Nuyt, anne.monique.nuyt@recherche-ste-justine.qc.ca

CMAJ 2012; DOI:10.1503/cmaj.120143



www.stat.gouv.qc.ca

Appariement des données

Définition :

- ❖ Consiste à **coupler** chaque enregistrement particulier d'un fichier donné avec l'enregistrement représentant la même unité dans un second fichier.

Sources des données :

- ❖ Enquêtes;
- ❖ Registre des événements démographiques (naissances vivantes, mortinaissances, décès);
- ❖ Banques de données administratives (RAMQ, MSSS, etc.);
- ❖ Banques de données d'autres sources (chercheurs).

Institut de la statistique
Québec

9

www.stat.gouv.qc.ca

Appariement des données

Possibilité d'apparier de façon :

- ❖ **Déterministe** : basé sur la comparaison exacte d'une combinaison de variables (numéro d'assurance maladie, nom, prénom, date de naissance, etc.);
- ❖ **Probabiliste** : basé sur une méthode qui permet de tirer profit du pouvoir discriminant de chacune des variables.

Avantage de l'appariement :

- ❖ Réutilisation de l'information existante provenant de diverses sources de données afin de répondre à des questions de recherche nécessitant le suivi d'un grand nombre d'individus dans le temps.

Institut de la statistique Québec

10



Conclusion (3)

- Banques de données administratives
 - Sources importantes d'information pour la santé de la population
 - Importance de la qualité des données et du maintien de cette qualité dans le temps
 - Importance de la pertinence des variables
 - Accès aux chercheurs




Remerciements

- Institut de la Statistique du Québec

- Ariane Boivin, Ph. D.
- Zhong-Cheng Luo, M.D., Ph. D.
- François Audibert, M.D.
- Benoit Mâsse, Ph. D.
- Francine Lefebvre, M.D.
- Réjean Tessier, Ph. D.



www.stat.gouv.qc.ca



Merci de votre attention!

Isabelle Leroux Analyste-conseil Direction des services informationnels et technologiques Institut de la statistique du Québec 418-691-2402, poste 3028	Anne Monique Nuyt, MD Professeur titulaire Service de néonatalogie- Dpt de pédiatrie CHU Ste-Justine Université de Montréal 514-345-4931, poste 3971
------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------


Pour plus d'information :

Site Internet anne.monique.nuyt@umontreal.ca

<http://www.stat.gouv.qc.ca/sad/>

Pour nous joindre par courriel :

SAD@stat.gouv.qc.ca
1 800-463-4090



Centre de
Recherche du
CHU Sainte-Justine
*Le centre hospitalier
universitaire mère-enfant*

Institut
de la statistique
Québec

14

