

---

# Identifying potentially eligible subjects for research: paper-based logs versus the hospital administrative database

---

L. A. Magee, MD (1,2,3,4); K. Massey MSc (2); P. von Dadelszen, DPhil (2,3,4); M. Fazio, BSc (2); B. Payne, BSc (2); R. Liston, MB ChB (2,4)

---

This article has been peer reviewed.

## Abstract

**Introduction:** The Canadian Perinatal Network (CPN) is a national database focused on threatened very pre-term birth. Women with one or more conditions most commonly associated with very pre-term birth are included if admitted to a participating tertiary perinatal unit at 22 weeks and 0 days to 28 weeks and 6 days.

**Methods:** At BC Women's Hospital and Health Centre, we compared traditional paper-based ward logs and a search of the Canadian Institute for Health Information (CIHI) electronic database of inpatient discharges to identify patients.

**Results:** The study identified 244 women potentially eligible for inclusion in the CPN admitted between April and December 2007. Of the 155 eligible women entered into the CPN database, each method identified a similar number of unique records (142 and 147) not ascertained by the other: 10 (6.4%) by CIHI search and 5 (3.2%) by ward log review. However, CIHI search achieved these results after reviewing fewer records (206 vs. 223) in less time (0.67 vs. 13.6 hours for ward logs).

**Conclusion:** Either method is appropriate for identification of potential research subjects using gestational age criteria. Although electronic methods are less time-consuming, they cannot be performed until after the patient is discharged and records and charts are reviewed. Each method's advantages and disadvantages will dictate use for a specific project.

---

**Keywords:** *subject identification, audit, health survey, hospital records, health records, database*

---

## Introduction

All clinical research studies begin with identifying potentially eligible subjects. Subjects can be identified by reviewing paper-based hospital or other health records designed for clinical purposes and by querying electronic patient databases used for administrative and/or clinical purposes.

The Canadian Perinatal Network (CPN) is a national perinatal database of women with threatened very pre-term birth at 22<sup>0</sup> to 28<sup>6</sup> weeks' gestation (22 weeks and 0 days to 28 weeks and 6 days) admitted to Canadian tertiary perinatal units. CPN began collecting data in August 2005, and by August 2009 involved 14 of Canada's 23 tertiary perinatal units. CPN-eligible patients must be identified for inclusion

based on their presentation to one of the participating units with one of the major causes of threatened very pre-term birth. CPN is a continuous quality improvement project with all data collection performed from patient health records.

Within our collaborating centres, the question arose as to the best method of identifying potentially eligible women for inclusion in CPN, since different methods are in use in different centres. These are either traditional paper-based admission records and ward logs or the Canadian Institute for Health Information (CIHI) electronic database of inpatient discharges. As a result, we sought to compare the two methods at the largest CPN centre, BC Women's Hospital and Health Centre in Vancouver.

## Methods

By August 1, 2009, CPN was enrolling patients from 14 of Canada's 23 tertiary perinatal units from centres in British Columbia (n = 2 centres), the Prairie provinces (n = 4), Ontario (n = 3), Quebec (n = 3) and the Atlantic provinces (n = 2). CPN was approved in each centre as a continuous quality improvement project.

Women are included in the CPN if they are admitted to a participating tertiary perinatal unit at 22<sup>0</sup> to 28<sup>6</sup> weeks with one or more of the conditions most commonly associated with very pre-term birth: spontaneous

---

### Author references:

1. Department of Medicine, Child and Family Research Institute, Vancouver, British Columbia, Canada
2. Department of Obstetrics and Gynaecology, Child and Family Research Institute, Vancouver, British Columbia, Canada
3. School of Population and Public Health, Child and Family Research Institute, Vancouver, British Columbia, Canada
4. Canadian Perinatal Network (CPN) Collaborative Group, Vancouver, British Columbia, Canada

Correspondence: Laura A. Magee, BC Women's Hospital and Health Centre, 4500 Oak Street, Room D213, Vancouver, BC V6H 3N1; Tel.: (604) 875-2424 ext. 6012; Fax: (604) 875-2961; Email: LMagee@cw.bc.ca

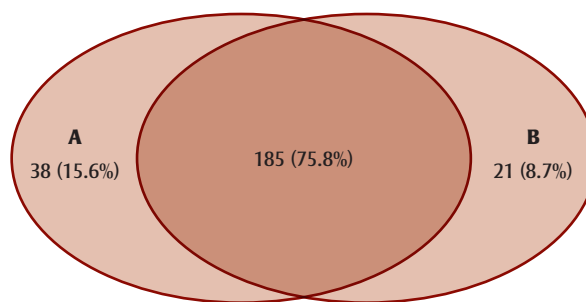
pre-term labour with contractions, incompetent cervix, prolapsing membranes, pre-term pre-labour rupture of membranes (PPROM), gestational hypertension, intrauterine growth restriction (IUGR), and/or antepartum hemorrhage (APH).\*

Women are excluded from the CPN if they are monitored for less than 24 hours in a triage area or obstetrical day unit and then sent home without being admitted to hospital. If the woman is admitted to hospital but later discharged, all subsequent re-admissions are recorded in CPN, up to and including her delivery.

Data abstractors identify women in one of two ways. First, delivery suite and antenatal ward records contain a patient log that includes the patient's name, gestational age, admission location, admission date, and depending on the location the hospital record number (delivery suite only). These paper-based data are collected and manually recorded by the nursing staff to administer clinical care and patient flow throughout the hospital. They are handwritten, often in pencil, and sometimes the names are erased or misspelled. In nine CPN centres, these logs are reviewed either in real-time or retrospectively by the CPN data abstractor, using the gestational age criteria of 22<sup>0</sup> to 28<sup>6</sup> weeks. In the five other CPN centres, a data abstractor requests a search of the centre's CIHI data through decision support staff; the query involves gestational age criteria of 22<sup>0</sup> to 28<sup>6</sup> weeks alone because admission diagnoses (as opposed to the final diagnoses made after delivery) are not recorded. The search output yields the mother's hospital identification number, gestational age, admission date, location of inpatient care and chief medical condition determined after delivery. Both approaches yield potentially eligible patients whose medical charts are then reviewed by the CPN site data abstractor who further defines eligibility and, when this is confirmed, abstracts the relevant patient data into the CPN database.

Data collection for CPN started at BC Women's Hospital in August 2005. Initially, the paper-based system of ward logs was used to identify potential

**FIGURE 1**  
Identification of potentially CPN-eligible women entered into the database (N=244)



(A) 223 cases identified by paper-based log searches; (B) 206 cases identified by querying the CIHI database.

**Abbreviations:** CIHI, Canadian Institute for Health Information; CPN, Canadian Perinatal Network.

subjects, and copies of these records were kept on file until the medical records of all potentially eligible women had been reviewed. Ward logs were obtained from BC Women's labour and delivery suite, antepartum unit and four postpartum units. In January 2008, subjects started to be identified through an electronic search of the CIHI database for gestational ages 22<sup>0</sup> to 28<sup>6</sup> weeks. This initial search was done back to January 2007, creating an overlap in identification methods for the period between April 1, 2007, and December 31, 2007 (the period for which ward logs had still been retained). For this period of overlap, the data abstractor reviewed the list of potential eligible subjects identified by CIHI to identify other potentially eligible women who may have been missed by the patient logs.

In July 2009, the patient ward logs available for the period between April 1, 2007, and December 31, 2007, were compared with a corresponding CIHI database search of locally retained data sent to CIHI by the hospital for gestational ages 22<sup>0</sup> to 28<sup>6</sup> weeks by a single reviewer who was not aware of which women were actually eligible and enrolled in CPN. We sought to determine the accuracy of paper-based versus electronic search methods of subject identification, as well as the time requirements for each approach, with results expressed descriptively as N (%).

## Results

From April 1, 2007, until December 1, 2007, a total of 244 women were identified as potentially eligible for enrolment in CPN at BC Women's Hospital based on gestational age criteria (22<sup>0</sup> – 28<sup>6</sup> weeks). Figure 1 shows that 185 (75.8%) women were identified by both the paper-based ward log review and the CIHI database output. Each method also identified a small number of women who were not identified by the other method: 38/244 (15.6%) for ward logs and 21/244 (8.7%) for CIHI. Review of the ward logs revealed missing or incorrect information such as surname spelling errors (confirmed on subsequent chart reviews) in 11/223 (4.9%) records. This prevented the data abstractor from further tracking the patient if no other identifiers were present, such as a hospital identification number, which is recorded routinely only by the delivery suite at the BC Women's Hospital.

From April 1, 2007, until December 1, 2007, records for 155 women were entered into the CPN database (at the BC Women's Hospital site) after manual review of their health records confirmed their eligibility. Figure 2 shows that 137/155 (88.4%) were identified by both the paper-based and electronic database search methods. Similar numbers of women were identified by only one of the two methods: ward logs captured 142/155 of the eligible women (91.6%) including 5 women (3.2%) who

\* For definitions of indicator conditions and maternal and perinatal outcomes see [http://www.cpn-rpc.org/doc/Appendix1\\_JOGC\\_20100726.pdf](http://www.cpn-rpc.org/doc/Appendix1_JOGC_20100726.pdf).

the CIHI had missed, and the CIHI data identified 147 of the eligible women (94.8%) including 10 women (6.4%) who were missed during the ward log review. There were also three women (1.9%) who were included in CPN but were neither identified by ward log review nor by CIHI search; these must have been identified by other means, such as word of mouth.

It took 13.6 hours to review the paper-based ward logs (i.e. 8 hours to search through the labour and delivery suite logs, 3.8 hours for antepartum unit logs and 1.8 hours for postpartum ward logs). These records had already been photocopied and were assembled and on file, so the actual time required to use the paper-based ward logs for patient selection would be longer. In contrast, the Decision Support Analyst took 0.67 hours to perform an electronic search of the Hospital CIHI data (0.50 hours to set up the initial query and 0.17 hours to run the initial query and each of any subsequent queries and forward the information to the CPN data abstractor).

## Discussion

The CPN uses two major methods to identify patients: review of paper-based ward logs and electronic search of the Hospital CIHI administrative databases, using gestational age-based criteria. The results of our analysis at BC Women's Hospital and Health Centre, the largest CPN site, showed that both of these approaches identifies the vast majority

(88%) of eligible women. The CIHI search identified a further 6.4% of unique records that were not identified by the ward logs, while a search of ward logs identified a further 3.2% that were not identified by the CIHI search. The CIHI search took substantially less time (0.67 hours, which included the initial query set-up, versus at least 13.6 hours for the paper-based ward logs because this estimate did not reflect the time taken to collect and photocopy the ward logs).

Review of ward logs has the advantage that it can be done daily, which permits prospective identification of patients. Conversely, a limitation of ward logs is missing or incorrect data (e.g. incorrect spelling of family name, wrong gestational age), which is not surprising as these logs are not intended for research purposes but to plan nursing assignments and manage admissions and discharges. Ward logs may also be difficult to double-check as a result of illegible hand-writing; it is possible that this is the reason for the three entries in the CPN database that were neither identified by CIHI search nor by ward log search. Such an omission may occur within a single shift, when a name is written in pencil and then is removed again, leaving no permanent record. Further, collecting these records, particularly from multiple locations, is time-consuming.

Electronic search of hospital administrative data has the advantage of being efficient and reproducible. It can perform more complex searches using structured query

language (depending on the clinical question and available data fields).<sup>1</sup> It also has the potential to search actual clinical records with increasing use of an electronic health record based on standardized language.<sup>2</sup> A limitation is miscoding, which is least likely to occur when basic terms (like "gestational age") are used.<sup>3</sup> The major limitation of this approach is that it cannot be done prospectively or in near real-time. Data are available only after the patient has been discharged and charts have been reviewed and abstracted in the Health Records department, which may take months in some institutions. As such, this method would not be feasible for researchers who need to identify women at or shortly after admission to hospital.

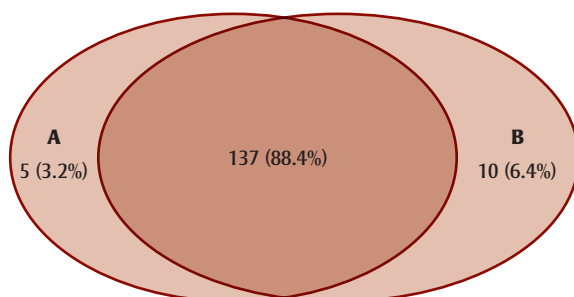
## Limitations

There are potential limitations to our study. The abstractor who performed this comparison of ascertainment methods was not biased by the initial eligibility assessment, as he did not do the initial review and CPN data entry; however, we were not able to measure inter-rater reliability. Our project relates to using gestational age criteria because neither ward logs nor CIHI data have additional admission diagnoses. However, the accuracy of using CIHI data might be different if additional relevant CIHI terms were available for another project. On the other hand, ward logs are very basic with regard to the information that they contain. Also, additional criteria for review of ward logs and/or CIHI searches may have yielded different results.

## Conclusion

Our study suggests that using gestational age-based criteria and either paper-based ward logs or electronic searches of hospital CIHI administrative database are both reasonably accurate methods of identifying potential subjects for clinical audit. Each method has its advantages and disadvantages, but database approaches are far less time-consuming, though they cannot be performed in or near real-time but only until after the patient has been discharged and information is abstracted from the ward logs.

**FIGURE 2**  
CPN-eligible women entered into the database (N=155)



(A) 142 cases identified by paper-based log searches; (B) 147 cases identified by querying the CIHI database.

**Abbreviations:** CIHI, Canadian Institute for Health Information; CPN, Canadian Perinatal Network.

---

## Acknowledgements

This study was supported by a grant from the Canadian Institutes of Health Research.

Thank you to the CPN site coordinators and previous CPN National Coordinators, D. Chaplain and T. Morris.

## References

1. Jamison DC. Structured Query Language (SQL) fundamentals. *Curr Protoc Bioinformatics*. 2003 Feb;Chapter 9:Unit9.2.
2. Massey KA, Ansermino JM, von Dadelszen P, Morris TJ, Liston RM, Magee LA. What is SNOMED CT and why should the ISSHP care? *Hypertens Pregnancy*. 2009 Feb;28(1):119-21.
3. Schulz EB, Barrett JW, Price C. Read Code quality assurance: from simple syntax to semantic stability. *J Am Med Inform Assoc*. 1998 Jul;5(4):337-46.