
Validation of perinatal data in the Discharge Abstract Database of the Canadian Institute for Health Information

K. S. Joseph, MD, PhD (1) and J. Fahey, MMath (2) for the Canadian Perinatal Surveillance System

Abstract

We compared perinatal information submitted to the Canadian Institute for Health Information (CIHI) hospitalization database with information submitted to the Nova Scotia Atlee Perinatal Database (NSAPD) in order to assess the accuracy of the CIHI data. Procedures such as Caesarean delivery were coded accurately (i.e. sensitivity of 99.8%; specificity of 98.7%). Postpartum hemorrhage, induction of labour and severe intraventricular hemorrhage also had sensitivity and specificity rates above 85% and 95%, respectively. Some diagnoses, defined differently in the two databases, were less accurately coded, e.g. respiratory distress syndrome (RDS) had a sensitivity of 50.9% and a specificity of 99.8%. Restriction to more severe forms of the disease improved accuracy, e.g. restriction of RDS to severe RDS in the NSAPD and identification of severe RDS in the CIHI database, using codes for RDS and intubation, resulted in a sensitivity of 100% and a specificity of 99.6%. Our study supports the use of CIHI data for national surveillance of perinatal morbidity, with the caveat that an understanding of clinical practice and sensitivity analyses to identify robust findings be used to facilitate inference.

Key words: *perinatal, surveillance, database, maternal, infant, morbidity, respiratory distress, discharge abstract data*

Introduction

Perinatal health surveillance in Canada relies on data from various sources, including vital statistics databases and hospital discharge databases.¹⁻³ Although vital statistics data on births and fetal and infant deaths remain an important source for documenting temporal trends and regional variations in perinatal health, the decline in mortality rates in recent decades has shifted the focus of perinatal surveillance increasingly towards monitoring trends and patterns in serious morbidity. This is particularly true with regard to serious maternal morbidity⁴ and serious neonatal morbidity.^{5,6}

The quality of hospitalization data from the Canadian Institute for Health Information (CIHI) Discharge Abstract Database (DAD) is an important concern. Although previous studies have concluded that this data source is suitable for surveillance purposes,⁷ a recent medical chart re-abstraction study, commissioned by the Canadian Perinatal Surveillance System (CPSS) and carried out by CIHI, showed variable quality with regard to several indicators of perinatal health.⁸ The study on hospital discharges in 1999/2000 included hospitals selected randomly after stratifying by geography and size and type of hospital. The charts of 385 newborns and 872 mothers were re-abstracted by CIHI classification specialists and compared

with information in the DAD. False positive rates for Caesarean, vacuum and forceps deliveries, preterm labour and episiotomies were < 1%; however, the false positive rates for other indicators were high (e.g. medical induction of labour, 12.8%; third degree perineal lacerations, 40.3%). Similarly, false negative rates were low for some indicators (e.g. < 1% for Caesarean deliveries, third degree lacerations, respiratory distress syndrome (RDS), pre-eclampsia / eclampsia), but high for other indicators (e.g. fetal asphyxia/fetal distress, 23.6%; medical induction of labour, 38.3%; rare neonatal conditions, 41.3%; rare congenital conditions, 53.9%).

Two concerns were voiced within the Canadian Perinatal Surveillance System (CPSS) Maternal Health and Fetal and Infant Health Study Groups who routinely use CIHI data for perinatal surveillance purposes.⁹⁻¹³ First, detailed analyses of phenomena such as amniotic fluid embolism,¹² postpartum hemorrhage¹³ and other conditions⁹⁻¹¹ have shown patterns that are congruent with clinical expectation, and suggest a higher level of data quality than that observed in the re-abstraction study. Another reason for questioning the high error rates in the re-abstraction study was related to the technical aspects of the stratified sampling (and the weighted calculation of population rates of false positive and false negative errors). Several error rates were relatively small in the study sample, but were substantially inflated after the population weights were applied. The most serious inflation was observed for third degree perineal

Author References

1 Perinatal Epidemiology Research Unit, Departments of Obstetrics & Gynaecology and Pediatrics, Dalhousie University and the IWK Health Centre, Halifax, NS, Canada.

2 Reproductive Care Program of Nova Scotia, Halifax, NS, Canada.

Correspondence: K.S. Joseph, MD, Division of Neonatal-Perinatal Medicine, IWK Health Centre, 5980 University Avenue, Halifax, NS B3K 6R8, Tel.: 902-470-6652; Fax: 902-470-7190; E-mail: ksjoseph@dal.ca

lacerations (i.e. false positive rate in the sample, 5.6%; in the population, 40.3%); rare neonatal conditions (i.e. false negative rate in the sample, 1.6%; in the population, 41.3%); and rare congenital conditions (i.e. false negative rate in the sample, 3.1%; in the population, 53.9%).

Consequently, we decided to reassess the accuracy of the DAD information by comparing the perinatal data in the DAD with data in the NSAPD. The latter database, which is smaller and clinically focused, is believed to have a relatively high degree of accuracy with regard to diagnoses and procedures.

Methods

During a brief period in 2002, the perinatal data pertaining to newborns and mothers in Nova Scotia were simultaneously coded for both the DAD and the NSAPD. Although the coding rules for the two databases were different, the availability of data under the two independent systems for 6194 mothers and 6315 newborns permitted an evaluation of the DAD data for quality assessment purposes. The mothers and newborns included in this duplicate coding were not selected by gestational age or outcome, and represented all of the deliveries/ births that occurred during a specific period.

The NSAPD is a clinically focussed, population-based database that stores detailed information from antenatal and medical charts. The information is extracted by trained personnel using standardized forms. An ongoing data quality assurance program, which carries out periodic abstraction studies, has shown the database information is reliable. The database has been used to validate the vital statistics, birth-and-infant-death-linked files at Statistics Canada.^{14,15} Perinatal information for the DAD was also collected by trained personnel in Nova Scotia under the CIHI data abstraction rules. For the period in question, the data were coded using International Classification of Diseases Revision 10 (ICD-10-CA) codes for diagnostic information and the Canadian Classification for Health Interventions (CCI) codes for interventions/procedures.

We compared diagnoses and interventions/procedures of interest between the two databases, assuming the NSAPD represents the gold standard. Rates of sensitivity (i.e. proportion of true [NSAPD] positives identified as being positive by the DAD) and specificity (i.e. proportion of true [NSAPD] negatives identified as being negative by the DAD) were calculated along with exact binomial 95% confidence intervals (CI). An evaluation of gestational age estimates from the two sources was also carried out using agreement statistics (i.e. weighted kappa and the intraclass correlation coefficient). For this analysis, gestational age was grouped into clinically relevant prognostic categories routinely used by the Canadian Perinatal Surveillance System (< 20, 20 to 21, 22 to 23, 24 to 25, 26 to 27, 28 to 31, 32 to 33, 34 to 36, 37 to 41, 42 to 45 weeks and unavailable). The specific diagnoses and interventions/procedures identified for assessing the accuracy of the DAD data were based on clinical and public health relevance and on the definitional compatibility of the diagnoses and interventions/procedures in the two databases.

Results

According to the DAD, the rate of preterm delivery (i.e. proportion of women with information on gestational age who delivered prior to 37 completed weeks) was 9.1% (95% CI 8.4 to 9.9%), whereas this rate was 8.8% (95% CI 8.1 to 9.6%) according to the NSAPD. The rate of postterm delivery (i.e. delivery at or after 42 completed weeks) was 0.6% (95% CI 0.5 to 0.9%) and 2.1% (95% CI 1.8 to 2.5%) according to the DAD and the NSAPD, respectively. No gestational age information was stated for 54 women in the NSAPD and for 7 women in the DAD (i.e. 47 women with missing gestational age in the NSAPD had a gestational age between 37 and 41 weeks, according to the DAD). Of the 543 women who had a preterm delivery according to the NSAPD, 495 were coded as having delivered preterm according to the DAD (i.e. sensitivity of 91.2%). Of the 5597 women who delivered at term or postterm gestation according to the NSAPD, 5531 were coded as having

delivered at term or postterm gestation according to the DAD (i.e. specificity of 98.8%). A detailed examination of the data on preterm delivery showed that 64 (i.e. 97%) of the 66 women coded as having delivered preterm by the DAD (but at term/postterm by the NSAPD) were at 36 weeks gestation according to the DAD, i.e. a large proportion of the false positive errors were due to a minor one-week difference in gestational age. Similarly, 31 of 48 (i.e. 65%) of the women who delivered at preterm gestation according to the NSAPD, but at term or postterm gestation according to the DAD, were at 37 weeks of gestational age at delivery according to the DAD. The weighted kappa statistic assessing the agreement between gestational age from the DAD data and gestational age from the NSAPD data was 0.75 (i.e. 95% CI 0.72 to 0.78), and the intraclass correlation coefficient was 0.86 (i.e. 95% CI 0.83 to 0.88).

Table 1 shows the sensitivity and specificity rates for several maternal health indicators. Most indicators in the DAD showed a high degree of accuracy. Sensitivity rates of 85% to 90% were noted for blood transfusions, induced labour and any gestational hypertensive disorders. There was good agreement between the many specific types of induction procedures/agents coded in the DAD and induction of labour coded in the NSAPD. For example, the 29 women who, according to the DAD, had labour induced through a combination of routes and involving the use of an oxytocic agent, were also coded as having their labour induced according to the NSAPD. The main discrepancy between the two databases occurred for cases where the DAD coded labour as having been induced by artificial rupture of membranes, whereas the NSAPD coded many such cases (i.e. 45 of 191) as having artificial rupture of membranes, but not as having been induced. Since artificial rupture of membranes after the onset of labour does not constitute labour induction, this discrepancy probably reflects a coding error in the DAD (i.e. assuming the information in the NSAPD is correct). Discrepancies were also noted for the diagnosis of hypertension in pregnancy. This was expected, because of the varied subtypes of hypertensive disorders in

TABLE 1
Validity of maternal data from the Discharge Abstract Database of the Canadian Institute for Health Information, using data from the Nova Scotia Atlee Perinatal Database as the gold standard (based on 6194 mothers, Nova Scotia, 2002)

Indicator	Sensitivity (%)	95% CI	Specificity (%)	95% CI
Preterm delivery (< 37 weeks)	91.2	88.5 to 93.4	98.8	98.5 to 99.1
Postpartum hemorrhage	90.2	86.2 to 93.3	98.2	97.8 to 98.5
Blood transfusion	85.7	42.1 to 99.6	99.8	99.6 to 99.9
Induction of labour	89.2	87.7 to 90.6	96.9	96.4 to 97.4
Caesarean delivery	99.8	99.5 to 100.0	98.7	98.3 to 99.0
Perineal laceration to – 1st degree	91.7	89.7 to 93.3	97.9	97.4 to 98.4
– 2nd degree	97.7	96.8 to 98.3	99.1	98.7 to 99.4
– 3rd degree	97.1	92.7 to 99.2	99.9	99.8 to 100.0
– 4th degree	94.7	74.0 to 99.7	99.9	99.8 to 100.0
Hypertension - chronic	83.3	73.6 to 90.6	99.9	99.8 to 100.0
Gestational hypertension with proteinuria (vs. severe PIH or HELLP)	75.2	67.5 to 81.8	99.5	99.3 to 99.7
Any gestational hypertensive disorder (vs. mild or severe PIH or HELLP)	87.9	85.0 to 90.4	99.6	99.4 to 99.8

Preterm delivery refers to women delivering before 37 completed weeks of gestation.

PIH denotes pregnancy induced hypertension.

HELLP denotes hemolysis, elevated liver enzymes and low platelets count syndrome.

pregnancy and the different labelling/classification schemes used by the NSAPD and the ICD-10-CA system.

Table 2 presents the assessment of infant health indicators among the 6315 newborns. The false negative rate for bacterial sepsis was high (i.e. sensitivity of 38.4%). This rate improved when cases in the DAD were also identified using adult codes for sepsis in addition to newborn codes. The more serious grades of intraventricular hemorrhage and fracture of the clavicle were accurately coded in the DAD, while RDS had the same discrepancies as hypertensive disorders of pregnancy (i.e. RDS classification is highly detailed in the NSAPD and differs from the diagnostic entities in the ICD-10-CA system). Nevertheless, combining an ICD-10-CA code of RDS with a procedure code of intubation in the DAD resulted in virtual agreement with a diagnosis of severe RDS in the NSAPD. Severe RDS in the NSAPD refers to RDS requiring assisted ventilation. Fetal/birth asphyxia was essentially not coded in the DAD (i.e. sensitivity of 14.3%, Table 2).

Discussion

Our study confirmed that major procedures such as Caesarean delivery were coded accurately in the DAD of CIHI. It also showed that the information on more minor diagnoses (e.g. first to fourth degree perineal lacerations) and more challenging diagnoses and procedures (e.g. induction of labour, which is easily confused with augmentation of labour) was also reasonably accurate. Similarly, gestational age—a difficult entity to capture accurately, given different methods of ascertaining gestational age—showed a relatively high degree of agreement between the two sources. The overall preterm birth rates were non-significantly higher in the DAD compared to the NSAPD, whereas postterm birth rates were significantly lower. Although these differences were relatively small, the direction of the differences suggests a greater influence of early ultrasound (and, hence, greater accuracy) on gestational age estimates in the DAD. In comparison with menstrual date-based estimation of gestational age, early ultrasound dating tends to slightly increase preterm birth rates and substantially lower postterm birth rates.¹⁶⁻¹⁸ The disagreements in the two databases arose mainly with regard to variably defined diagnostic entities, such

as bacterial sepsis and RDS. Nevertheless, identifying bacterial sepsis using both neonatal and adult codes for sepsis and confining RDS to “any respiratory distress” or to a severe form of RDS resulted in relatively accurate information for the DAD.

The serious discrepancy between the NSAPD and the DAD with regard to fetal/birth asphyxia is to be expected and is not a reflection of data inaccuracy in the DAD. Studies have shown that this diagnostic label has essentially disappeared¹⁹ from the clinical lexicon due to malpractice concerns, even though the clinical entity remains essentially unchanged in its frequency.²⁰ Thus, it is rarely captured in the DAD system, where coders identify cases only if the term “asphyxia” is documented in the medical chart, while the NSAPD continues to identify the condition based on its clinical components.

This study was carried out using medical records from 2002, a time when Nova Scotia first began to implement the ICD-10-CA coding in its hospitals. Since the ICD-10-CA system is substantially different from the previous version, some coding errors may have occurred in 2002 that would have been resolved with continued use. One possible example of this may be seen with regard

TABLE 2
Validity of neonatal data from the Discharge Abstract Database of the Canadian Institute for Health Information, using data from the Nova Scotia Atlee Perinatal Database as the gold standard (based on 6315 live births, Nova Scotia, 2002).

Indicator	Sensitivity (%)	95% CI	Specificity (%)	95% CI
Bacterial sepsis	38.4	28.1 to 49.5	99.7	99.5 to 99.8
Bacterial sepsis (adult/neonatal codes)	67.4	56.5 to 77.2	99.6	99.4 to 99.8
Intraventricular hemorrhage, grade 3, 4	88.9	51.8 to 99.7	100.0	99.9 to 100.0
Fracture of clavicle	91.7	61.5 to 99.8	100.0	99.3 to 100.0
Respiratory distress – any (vs. any)*	94.2	90.8 to 96.6	96.6	96.1 to 97.1
– RDS (vs. RDS)	50.9	43.1 to 58.6	99.8	99.7 to 99.9
– RDS (vs. severe RDS)	96.3	89.6 to 99.2	99.6	99.4 to 99.8
– any RDS + intubation (vs. severe RDS)	100.0	95.5 to 100.0	99.6	99.4 to 99.8
Fetal/birth asphyxia	14.3	6.4 to 26.2	99.3	99.1 to 99.5

* Any code vs. any code refers to an evaluation of any respiratory distress code in CIHI vs. any respiratory distress code in the NSAPD.

RDS vs. RDS refers to a respiratory distress syndrome code in CIHI vs. a respiratory distress syndrome code in the NSAPD.

RDS vs. severe RDS refers to a respiratory distress syndrome code in CIHI vs. a severe respiratory distress syndrome code in the NSAPD.

Any RDS + intubation refers to an evaluation of a respiratory distress syndrome code in CIHI plus an intubation code vs. a severe respiratory distress syndrome code in the NSAPD.

to the bacterial sepsis of the newborn codes, which had a poor sensitivity rate arising at least partly because adult sepsis codes were used. Presumably, such errors would have been corrected as familiarity with the ICD-10-CA system increased.

The improved diagnostic accuracy of diseases, such as RDS in the DAD with the restriction to more severe forms of the disease, suggests that researchers using these large hospitalization databases should routinely carry out sensitivity analyses to assess the robustness of their findings. Familiarity with the clinical culture and changes in clinician habits (e.g. knowledge of the declining use of terms, such as birth asphyxia)^{19,20} is a critical factor in appropriately interpreting patterns in large databases. A multi-disciplinary approach to research using such databases is probably the most appropriate approach.

One limitation of our study arose because, in most instances in Nova Scotia, the same health-records personnel coded medical charts for both the DAD and the NSAPD. This may mean that agreement between the two systems is higher than would be expected if the two systems were fully independent. On the other hand, both systems have detailed rules regarding coding (including criteria for diagnosis and specified parts of

the medical chart from which information is to be extracted). These coding rules are evident in the discrepancies in diagnoses for which standard diagnostic criteria were unavailable.

Other potential limitations of our study include the assumption that the NSAPD, a smaller, clinically focused database with a data quality assurance program, is more accurate than the DAD. This assumption, though tenable, is unlikely to hold with regard to all of the information in the two databases. For example, the CIHI database collects information on blood and blood products in much greater detail than the NSAPD does. Furthermore, the NSAPD algorithm for determining gestational age relies solely on menstrual dates and a pediatric examination of the newborn infant. Conversely, gestational age in the DAD could represent a better estimate, because it also incorporates early ultrasound information.¹⁶⁻¹⁸

Finally, our assessment was limited to one Canadian province and, hence, may not be generally applicable to other provinces/territories, or even to Nova Scotia at a different time. Nevertheless, the relatively high level of accuracy observed in our study is encouraging and more in line with the expectations of CPSS investigators

who have worked with DAD data⁹⁻¹³ than the results of the abstraction-re-abstraction study.⁸

In summary, our study compared information in the DAD with that from a smaller clinically focused database, showing that the DAD information was accurate for many of the diagnoses/procedures examined. Furthermore, less accurate diagnoses, typically observed in the case of variably defined clinical entities, can be improved using combined codes and a restriction to more severe forms of the disease. This study therefore supports the use of the data in the CIHI DAD for national perinatal surveillance and research, with the caveat that appropriate inference rests on an understanding of clinical practice and the use of sensitivity analyses to identify robust findings.

Acknowledgements

This study was funded by the Maternal and Infant Health Section, Health Surveillance and Epidemiology Division of the Public Health Agency of Canada. We are grateful to the Reproductive Care Program of Nova Scotia for access to the data.

References

1. Health Canada. Canadian perinatal health report. Ottawa: Minister of Public Works and Government Services Canada; 2000.
2. Health Canada. Canadian perinatal health report. Ottawa: Minister of Public Works and Government Services Canada; 2003.
3. Public Health Agency of Canada. Canadian perinatal health report. 2008 ed. Ottawa: Minister of Public Works and Government Services Canada; 2008.
4. Wen SW, Huang L, Liston R, et al. Severe maternal morbidity in Canada, 1991–2001. *CMAJ*. 2005;173(7):759–64.
5. Wilson-Costello D, Friedman H, Minich N, Fanaroff AA, Hack M. Improved survival rates with increased neurodevelopmental disability for extremely low birth weight infants in the 1990s. *Pediatrics*. 2005;115:997–1003.
6. Wilson-Costello D, Friedman H, Minich N, Siner B, Taylor G, Schluchter M, Hack M. Improved neurodevelopmental outcomes for extremely low birth weight infants in 2000–2002. *Pediatrics*. 2007;119:37–45.
7. Wen SW, Liu S, Marcoux S, Fowler D. Uses and limitations of routine hospital admission/separation records for perinatal surveillance. *Chronic Dis Can*. 1997;18:113–9.
8. Health Canada. An evaluation of the quality of obstetric/neonatal discharge abstract data by reabstraction of medical charts. Ottawa: Health Canada; 2003.
9. Liu S, Heaman M, Kramer MS, Demissie K, Wen SW, Marcoux S. Length of hospital stay, obstetric conditions at childbirth, and maternal readmission: a population-based cohort study. *Am J Obstet Gynecol*. 2002;187:681–7.
10. Liu S, Heaman M, Joseph KS, et al. Risk of maternal postpartum readmission associated with mode of delivery. *Obstet Gynecol*. 2005;105:836–42.
11. Liu S, Liston RM, Joseph KS, Heaman M, Sauve R, Kramer MS. Maternal mortality and severe morbidity associated with low-risk planned cesarean delivery versus planned vaginal delivery at term. *CMAJ*. 2007;176:455–60.
12. Kramer MS, Rouleau J, Baskett TF, Joseph KS. Amniotic-fluid embolism and medical induction of labour: a retrospective, population-based cohort study. *Lancet*. 2006;368:1444–8.
13. Joseph KS, Rouleau J, Kramer MS, Young DC, Liston RM, Baskett TF. Investigation of an increase in postpartum haemorrhage in Canada. *BJOG*. 2007;114:751–9.
14. Fair M, Cyr M, Allen AC, Wen SW, Guyon G, MacDonald RC. An assessment of the validity of a computer system for probabilistic record linkage of birth and infant death records in Canada. *Chronic Dis Can*. 2000;21:8–13.
15. Fair M, Cyr M, Allen AC, Wen SW, Guyon G, Macdonald RC. Validation study for a record linkage of births and infant deaths in Canada. Ottawa: Statistics Canada; 1999.
16. Kramer MS, McLean FH, Boyd ME, Usher RH. The validity of gestational age estimation by menstrual dating in term, preterm, and postterm gestations. *JAMA*. 1988;260:3306–8.
17. Goldenberg RL, Davis RO, Cutter GR, Hoffman HJ, Brumfield CG, Foster JM. Prematurity, postdates, and growth retardation: the influence of use of ultrasonography on reported gestational age. *Am J Obstet Gynecol*. 1989;160:462–70.
18. Joseph KS, Huang L, Liu S, Ananth CV, Allen AC, Sauve R, Kramer MS. Reconciling the high rates of preterm and postterm birth in the United States. *Obstet Gynecol*. 2007;109:813–22.
19. Wu YW, Backstrand KH, Zhao S, Fullerton HJ, Johnston SC. Declining diagnosis of birth asphyxia in California: 1991–2000. *Pediatrics*. 2004;114:1584–90.
20. Dzakpasu S, Joseph KS, Huang L, Allen A, Sauve R, Young D. Decline in birth asphyxia in Canada: fact or artefact. *Pediatrics*. 2009;123:e668–72.

