
Estimating gestational age at birth: a population-based derivation-validation study

M. L. Urquia, PhD (1,2); T. A. Stukel, PhD (2,3); K. Fung, MSc (2); R. H. Glazier, MD (1,2,3,4,5); J. G. Ray, MD (2,3,6)

Abstract

Introduction: Information on newborn gestational age (GA) is essential in research on perinatal and infant health, but it is not always available from administrative databases. We developed and validated a GA prediction model for singleton births for use in epidemiological studies.

Methods: Derivation of estimated GA was calculated based on 130 328 newborn infants born in Ontario hospitals between 2007 and 2009, using linear regression analysis, with several infant and maternal characteristics as the predictor (independent) variables. The model was validated in a separate sample of 130 329 newborns.

Results: The discriminative ability of the linear model based on infant birth weight and sex was reasonably approximate for infants born before the 37th week of gestation ($r^2 = 0.67$; 95% CI: 0.65–0.68), but not for term births (37–42 weeks; $r^2 = 0.12$; 95% CI: 0.12–0.13). Adding other infant and maternal characteristics did not improve the model discrimination.

Conclusion: Newborn gestational age before 37 weeks can be reasonably approximated using locally available data on birth weight and sex.

Keywords: *gestational age, birth, neonate, infant health, derivation, validation, prediction, administrative datasets, Ontario*

Introduction

Gestation starts on the day of conception and ends at birth, but it is typically measured from the first day of the last menstrual period. Gestational age (GA) is a major predictor of perinatal mortality and morbidity;¹ it is important for dating for prenatal genetic screening² and for the timing of fetal exposure to teratogens.^{3,4} It is also needed to correctly determine if an infant is small or large for GA, both for clinical practice and epidemiological research.²

In countries where antenatal maternal care is scarce, the collection of basic newborn statistics may be hampered by a lack of information on GA. On the other hand, in industrialized nations, GA is often not recorded in administrative health databases.^{3–5} Since all permanent residents of Canada receive universal health care, including prenatal, peripartum and newborn care, the Discharge Abstract Database of the Canadian Institute of Health Information (CIHI-DAD), an administrative database, has been recognized as an excellent source for population-based estimates for

perinatal research;^{6,7} however, prior to fiscal year 2002/03, CIHI-DAD did not collect data on GA at birth in Ontario,⁸ which could pose problems for some perinatal outcomes research.

The aim of this study is to develop and validate a GA prediction model for singleton births for use in epidemiological studies.

Methods

General design

We used a derivation-validation analytical method to estimate GA based on commonly available perinatal data. We completed a large population-based study of all singleton infants born in Ontario hospitals in 2007/08 and 2008/09, the period during which GA at birth was fully recorded by CIHI-DAD. The derivation cohort consisted of a randomly selected sample of 50% of all live births in this same period. This cohort was used to generate a predictive model based on infant characteristics. The other 50% of births formed the validation cohort, to test the derivation model's prediction of GA at birth. Simulation studies have shown that split-sample validation is a reasonable approach when the overall sample size is very large, as in our study ($N = 260\,657$).⁹

We excluded all stillbirths and multiple births from our sample. To minimize the influence of potential data errors and outliers, we also excluded infants born at or less than 23 completed weeks gestation or at or more than 43 completed weeks gestation; those with clinically implausible combinations of birth weight and GA;¹⁰ those who

Author references

1. Centre for Research on Inner City Health, The Keenan Research Centre of the Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Ontario, Canada
 2. Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada
 3. Department of Health Policy Management and Evaluation, University of Toronto, Toronto, Ontario, Canada
 4. Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada
 5. Department of Family and Community Medicine, University of Toronto, Toronto, Ontario, Canada.
 6. Department of Obstetrics and Gynecology, University of Toronto, Toronto, Ontario, Canada
- Correspondence:** Marcelo Luis Urquia, Centre for Research on Inner City Health, The Keenan Research Centre of the Li Ka Shing Knowledge Institute, St. Michael's Hospital, 30 Bond Street, Toronto ON M5B 1W8; Tel.: (416) 864 6060 x 77340; Fax: (416) 864 5558; Email: marcelo.urquia@utoronto.ca

stayed in hospital for more than 90 days; those whose GA, birth weight or sex was not recorded; those born to mothers aged less than 16 years or over 50 years at the time of delivery; and extreme outliers of the birth weight distribution identified as values located outside the inter-quartile range exceeding two times its distance.¹¹

Variables

In Ontario, GA is largely estimated by early ultrasound dating. Since 2002, hospital medical records departments have recorded GA based on the attending physician's best interpretation of all clinical data, usually presented on the antenatal record.^{12,13} This, along with the infant's sex and precisely measured birth weight, is recorded in the CIHI-DAD.¹⁴ We determined congenital anomalies and diseases of prematurity from the ICD-10-CA* codes¹⁵ entered in the 25 diagnostic fields in the hospital records (Table 1).

Analysis

Derivation of the estimate of GA involved two steps.¹⁶ Using the derivation cohort, we performed a series of linear regression analyses with completed GA (in weeks) as the dependent variable and several independent variables, chosen a priori, as listed in Table 2.

We first modelled GA using a restricted cubic spline function of birth weight with four degrees of freedom.¹⁷ We added infant sex, congenital and chromosomal anomalies and the diseases of prematurity (respiratory distress syndrome, neonatal cerebral leukomalacia or intraventricular hemorrhage, retinopathy of prematurity, necrotizing enterocolitis) to the basic model. The details of these variables are listed in Table 2.

We generated prediction models by multiplying the coefficients with each independent variable in the derivation models by the specific values that make up the profile of each individual in the validation cohort. We tested each prediction model using the validation cohort's *true* GA as the dependent variable and *estimated* GA as the independent variable, rounded to the nearest completed week. As a measure of model

discrimination, we computed the coefficient of determination (r^2) and its 95% confidence interval (CI). Models were validated for the entire birth cohort, and stratified by infant sex and by timing of birth (less than 37 weeks GA and equal or more than 37 weeks GA). The true versus estimated GA was plotted according to their respective frequency distributions (Figure 1).

We plotted the true positive rate of the derived model (i.e. the proportion of infants whose true GA is equal to the derived GA, is within 1 week of derived GA, or is within 2 weeks of derived GA) on a y-axis against the estimated GA on the x-axis (Figure 2).

All analyses were conducted using SAS version 9.1 (SAS Institute Inc., Cary, NC, US).

Results

There were 281 406 infant records in 2007/08 and 2008/09. After excluding stillbirths and multiple births and obvious outliers (7.4%), the final available dataset consisted of 260 657 singletons. Infant characteristics in both the derivation and validation cohorts were similar (Table 2).

The optimal model included a restricted cubic spline function of birth weight (in kilogram) as well as infant sex. The coefficient of determination (r^2) for this predictive model was 0.44 (95% CI: 0.43–0.45). Adding any congenital or chromosomal anomaly or diseases of prematurity, or stratifying by infant sex to the above model did not appreciably affect the coefficient of determination (Table 3).

Stratifying by timing of birth, the discriminative ability of the model was poor for infants delivered at term (37–42 weeks: $r^2 = 0.12$; 95% CI = 0.12–0.13), but much better for preterm births (24–36 weeks: $r^2 = 0.67$; 95% CI = 0.65–0.68) (Table 3). Adding admission to a neonatal intensive care unit, infant hospital length of stay, maternal preeclampsia or gestational hypertension and mode of delivery to the pre-term model did not further improve the coefficient of determination (data not shown).

Up to about 36 weeks gestation, there was high concordance in the distribution curves for true versus derived GA, after which there was marked discordance (Figure 1). At term, predicted GA does not estimate the true GA well, especially at 39 weeks, when most infants are born (Figure 1).

The GA model that included infant birth weight and sex had a positive predictive value of 34% at 28 ± 1 weeks, 67% at 28 ± 2 weeks, 47% at 32 ± 1 weeks, 74% at 32 ± 2 weeks, 60% at 37 ± 1 weeks and 85% at 37 ± 2 weeks gestation (Figure 2).

We repeated the validation using the entire dataset instead of the validation dataset and the results did not change (data not shown).

Discussion

In a large population-based derivation-validation study, infant birth weight and sex together provided a reasonable estimate of GA among infants born before 37 weeks, but not among term infants.

The addition of other newborn and maternal characteristics did not improve the coefficient of determination of our model among preterm infants. Others have noted similar results in the development of newborn birth weight curves.¹⁸

A parsimonious model based on infant birth weight and sex has some advantages in that both variables are captured and recorded in nearly all clinical encounters within both poorer and wealthier nations and also within large administrative datasets in which GA is not available. It is noteworthy that infant birth weight and sex are the two main variables used for the construction of population-based references of birth weight for GA.^{10,19,20} Therefore, in the absence of recorded GA, we recommend using information on infant birth weight and sex to approximate GA, and figures from local birth weight for GA charts, including the observed sex-specific 50th percentile of birth weight at each week of GA. Lower (5th, 10th) and upper (90th, 95th) percentiles of birth

* International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Canadian Enhancement

TABLE 1
ICD-10-CA codes used to determine congenital anomalies, diseases of prematurity, multiple births and stillbirths among singleton live-born newborns, Ontario, 2007–09

Variable	CIHI-DAD record source	ICD-10-CA
Any congenital or chromosomal anomaly	Infant	Q00-Q99
Diseases of prematurity	Infant	
Necrotizing enterocolitis		P77
Respiratory distress syndrome		P22
Neonatal cerebral leukomalacia or intraventricular hemorrhage		P91.2, P52
Retinopathy of prematurity		H35.1
Multiple gestation	Infant	Q89.4, Z38.3-Z38.8
Multiple gestation	Maternal	O30, O31, Z37.2-Z37.7, Z38.3-Z38.8, Z37.9.0
Intrauterine death	Infant	P95
Intrauterine death	Maternal	O36.4, Z37.1, Z37.4, Z37.7

Abbreviations: CIHI-DAD, Discharge Abstract Database of the Canadian Institute of Health Information; ICD-10-CA, International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Canadian Enhancement.

TABLE 2
Characteristics of singleton live-born newborns in the derivation cohort (n = 130 328) and validation cohort (n = 130 329), Ontario, 2007–09

Infant characteristics	Derivation cohort, n (%)	Validation cohort, n (%)
Male	66 551 (51.06)	66 898 (51.33)
Gestational age at birth		
Term, 37–42 weeks	122 723 (94.16) ^a	122 760 (94.19) ^b
Preterm, 24–36 weeks	7 605 (5.84)	7 569 (5.81)
Very preterm, 24–27 weeks	187 (0.14)	206 (0.16)
Mean birth weight \pm SD, grams	3 392 \pm 531	3 392 \pm 532
Birth weight ^c		
< 2500 grams	5 715 (4.39)	5 797 (4.45)
\geq 2500 grams	124 613 (95.61)	124 532 (95.55)
Congenital or chromosomal anomalies ^d	5 655 (4.34)	5 677 (4.36)
Diseases of prematurity ^{d,e}	7 587 (5.82)	7 771 (5.96)
Respiratory distress syndrome	7 474 (5.73)	7 681 (5.89)
Neonatal cerebral leukomalacia or intraventricular hemorrhage	206 (0.16)	207 (0.16)
Retinopathy of prematurity	111 (0.09)	112 (0.09)
Necrotizing enterocolitis	62 (0.05)	62 (0.05)

Abbreviations: ICD-10-CA, International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Canadian Enhancement; n, sample size; SD, standard deviation.

^aThe mean gestational age (\pm SD) at birth in this group was 39.2 (\pm 1.14) weeks.

^bThe mean gestational age (\pm SD) at birth in this group was 39.2 (\pm 1.15) weeks.

^cThe mean birth weight (\pm SD) was 3392 (\pm 531) grams for the derivation cohort and 3392 (\pm 532) grams for the validation cohort.

^dCongenital or chromosomal anomalies and diseases or prematurity determined from ICD-10-CA codes in hospital records.

^eMany newborns may have more than one disease of prematurity. Hence the percentages do not add up to 100.

weight could also be used to express biological variability in GA at a given birth weight.

The finding that GA and birth weight are poorly correlated after 36 weeks gestation is noteworthy given that about 94% of singleton infants are born at term. The poor prediction of GA at term is basically due to the large variability in birth weight as GA increases. For example, a recent Canadian birth weight chart for male newborns showed a minimum 1100-gram difference between the 10th and the 90th percentiles of birth weight at 37 to 41 weeks gestation.²⁰ The latter reflects a large amount of variability in birth weight within the “normal” range of birth weight. The better prediction of GA at earlier gestational periods is reflective of less biological variability. In addition, the birth weight slope is more linear and steeper at lower GAs than at term.²⁰

Limitations

This study has a number of limitations. First, we relied on ICD-10-CA codes within an administrative database in which infant measurements were not performed for the purpose of this study. Second, we only included singleton live-born infants, so our approach may not apply to multiple pregnancies. Unfortunately, population-based birth weight curves for multiple births are scarce.^{21,22} Third, the database did not contain information on other factors associated with length of gestation and newborn weight, such as parental ethnicity, maternal anthropometry and health behaviours during pregnancy, each of which may be used in the construction of customised newborn weight charts.^{23,24} Inclusion of these factors might improve our prediction model.^{25,26} Fourth, we based our analyses on the clinical estimate of GA (typically based on early ultrasound dating), which is known to differ from the estimate based on the date of last menstrual period.^{12,13} The latter has been found to overestimate preterm and postterm birth rates and present bimodal birth weight distributions between 28 and 34 weeks of gestation.^{20,25,27–29} Replication of our validation approach using the menstrual estimate of gestation as the “gold standard” may likely lead to poorer prediction. Finally, we caution others that our models were not designed to specifically estimate the GA of individual newborns.

FIGURE 1
Comparison of the predicted gestational age based on infant birth weight and sex (solid line) versus the true gestational age at birth (dashed line), validation data (n = 130 329)

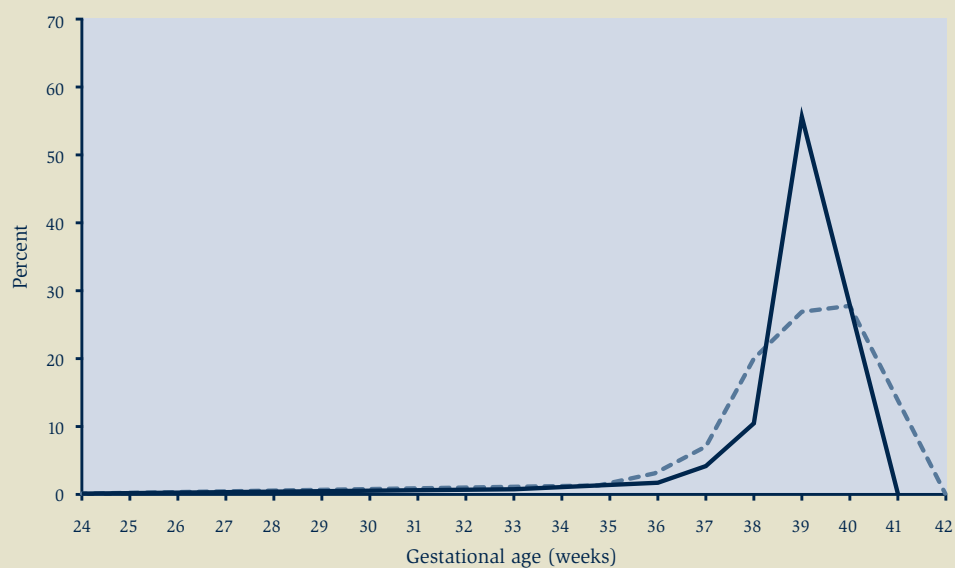


FIGURE 2
Agreement between derived gestational age and true gestational age among singleton live births in Ontario, 2007/08 to 2008/09. The curves represent the percentage of infants whose true gestational age is equal to the derived gestational age (lower), or is within ± 1 week (middle) and ± 2 weeks (upper)



TABLE 3
Discriminative ability of a model generated from the derivation cohort to predict known gestational age at birth in the validation cohort^a

Group	Model variables	Coefficient of determination, r^2 (95% CI) ^b
All (n = 130 329)	Birth weight and sex	0.44 (0.43–0.45)
	Birth weight, sex, congenital or chromosomal anomalies, and diseases of prematurity ^c	0.45 (0.44–0.46)
Sex		
Males (n = 66 898)	Birth weight	0.46 (0.44–0.47)
	Birth weight, congenital or chromosomal anomalies, and diseases of prematurity ^c	0.47 (0.45–0.48)
Females (n = 63 431)	Birth weight	0.43 (0.41–0.44)
	Birth weight, congenital or chromosomal anomalies, and diseases of prematurity ^c	0.44 (0.42–0.45)
Timing at birth		
Term, 37–42 weeks (n = 122 760)	Birth weight and sex	0.12 (0.12–0.13)
	Birth weight, sex, congenital or chromosomal anomalies, and diseases of prematurity ^c	0.13 (0.12–0.13)
Preterm, 24–36 weeks (n = 7569)	Birth weight and sex	0.67 (0.65–0.68)
	Birth weight, sex, congenital or chromosomal anomalies, and diseases of prematurity ^c	0.68 (0.67–0.70)

Abbreviations: ICD-10-CA, International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Canadian Enhancement; n, sample size.

^a Cohorts are singleton infants live born in Ontario in 2007/08–2008/09.

^b Birth weight is modeled as a restricted cubic spline with 4 degrees of freedom.

^c Respiratory distress syndrome, neonatal cerebral leukomalacia or periventricular hemorrhage, retinopathy of prematurity, necrotizing enterocolitis based on ICD-10-CA codes.

In conclusion, in the absence of information on actual GA, newborn GA can be reasonably approximated at the population level as a continuous variable up to 36 weeks gestation using birth weight and sex, although substantial uncertainty seems unavoidable, even after considering other predictors of GA.

Acknowledgements

This study was supported by the Institute for Clinical Evaluative Sciences (ICES), which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). The positions, results and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario MOHLTC is intended and nor should it be inferred.

References

- Cheng YW, Nicholson JM, Nakagawa S, Bruckner TA, Washington AE, Caughey AB. Perinatal outcomes in low-risk term pregnancies: do they differ by week of gestation? *Am J Obstet Gynecol.* 2008; 199(4):370-7.
- Behrman RE, Butler AS, editors; Committee on Understanding Premature Birth and Assuring Healthy Outcomes. *Preterm birth: causes, consequences, and prevention.* 1st ed. Washington (DC): National Academies Press; 2007.
- Andrade SE, Raebel MA, Morse AN, Davis RL, Chan KA, Finkelstein JA, et al. Use of prescription medications with a potential for fetal harm among pregnant women. *Pharmacoepidemiol Drug Saf.* 2006; 15(8):546-54.
- Hardy JR, Leaderer BP, Holford TR, Hall GC, Bracken MB. Safety of medications prescribed before and during early pregnancy in a cohort of 81,975 mothers from the UK General Practice Research Database. *Pharmacoepidemiol Drug Saf.* 2006;15(8):555-64.
- Toh S, Mitchell AA, Werler MM, Hernandez-Diaz S. Sensitivity and specificity of computerized algorithms to classify gestational periods in the absence of information on date of conception. *Am J Epidemiol.* 2008;167(6):633-40.
- Joseph KS, Fahey J; Canadian Perinatal Surveillance System. Validation of perinatal data in the Discharge Abstract Database of the Canadian Institute for Health Information. *Chronic Dis Can.* 2009;29(3):96-100.
- 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(3):113-9.
- Canadian Institute for Health Information. Data quality documentation: Discharge Abstract Database 2002–2003. Ottawa (ON): CIHI; 2005.
- Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol.* 2001;54(8):774-81.
- Alexander GR, Himes JH, Kaufman RB, Mor J, Kogan M. A United States national reference for fetal growth. *Obstet Gynecol.* 1996;87(2):163-8.
- Tukey JW. *Exploratory data analysis.* Reading (MA): Addison-Wesley; 1977.
- Wingate MS, Alexander GR, Buekens P, Vahratian A. Comparison of gestational age classifications: date of last menstrual period vs. clinical estimate. *Ann Epidemiol.* 2007;17(6):425-30.
- Qin C, Hsia J, Berg CJ. Variation between last-menstrual-period and clinical estimates of gestational age in vital records. *Am J Epidemiol.* 2008;167(6):646-52.

14. Canadian Institute for Health Information. DAD abstracting manual (for use with ICD-10-CA/CCI) 2006-2007. Ottawa (ON): CIHI; 2006.
15. Canadian Institute for Health Information. Final report: the Canadian enhancement of ICD-10 (International Statistical Classification of Diseases and Related Health Problems, Tenth Revision). Ottawa (ON): CIHI; 2001.
16. Sullivan LM, Massaro JM, D'Agostino RB Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med*. 2004;23(10):1631-60.
17. Durrleman S, Simon R. Flexible regression models with cubic splines. *Stat Med*. 1989;8(5):551-61.
18. Hutcheon JA, Zhang X, Cnattingius S, Kramer MS, Platt RW. Customised birthweight percentiles: does adjusting for maternal characteristics matter? *BJOG*. 2008;115(11):1397-404.
19. Bonellie S, Chalmers J, Gray R, Greer I, Jarvis S, Williams C. Centile charts for birthweight for gestational age for Scottish singleton births. *BMC Pregnancy Childbirth*. 2008;8:5.
20. Kramer MS, Platt RW, Wen SW, Joseph KS, Allen A, Abrahamowicz M, et al. A new and improved population-based Canadian reference for birth weight for gestational age. *Pediatrics*. 2001;108(2):E35.
21. Ananth CV, Vintzileos AM, Shen-Schwarz S, Smulian JC, Lai YL. Standards of birth weight in twin gestations stratified by placental chorionicity. *Obstet Gynecol*. 1998;91(6):917-24.
22. Glinianaia SV, Skjaerven R, Magnus P. Birthweight percentiles by gestational age in multiple births. A population-based study of Norwegian twins and triplets. *Acta Obstet Gynecol Scand*. 2000; 79(6):450-8.
23. Gardosi J, Francis A. A customized standard to assess fetal growth in a US population. *Am J Obstet Gynecol*. 2009;201(1):25.e1-7.
24. Gardosi J, Chang A, Kalyan B, Sahota D, Symonds EM. Customised antenatal growth charts. *Lancet*. 1992;339(8788):283-7.
25. Kierans WJ, Joseph KS, Luo ZC, Platt R, Wilkins R, Kramer MS. Does one size fit all? The case for ethnic-specific standards of fetal growth. *BMC Pregnancy Childbirth*. 2008;8:1.
26. Ray JG, Jiang D, Sgro M, Shah R, Singh G, Mamdani MM. Thresholds for small for gestational age among newborns of East Asian and South Asian ancestry. *J Obstet Gynaecol Can*. 2009;31(4):322-30.
27. Joseph KS, Huang L, Liu S, Ananth CV, Allen AC, Sauve R, et al. Reconciling the high rates of preterm and postterm birth in the United States. *Obstet Gynecol*. 2007;109(4):813-22.
28. Platt RW, Abrahamowicz M, Kramer MS, Joseph KS, Mery L, Blondel B, et al. Detecting and eliminating erroneous gestational ages: a normal mixture model. *Stat Med*. 2001;20(23):3491-503.
29. Tentoni S, Astolfi P, De PA, Zonta LA. Birthweight by gestational age in preterm babies according to a Gaussian mixture model. *BJOG*. 2004;111(1):31-7.