BREASTFEEDING MEDICINE
Volume 10, Number 10, 2015

DOI: 10.1089/bfm.2015.0126

Clinical Research

External Validation of Early Weight Loss Nomograms for Exclusively Breastfed Newborns

Eric W. Schaefer,[1] Valerie J. Flaherman,[2,3] Michael W. Kuzniewicz,[2,4]
Sherian X. Li,[4] Eileen M. Walsh,[4] and Ian M. Paul[1,5]

Abstract

Introduction: Nomograms that show hour-by-hour percentiles of weight loss during the birth hospitalization were recently developed to aid clinical care of breastfeeding newborns. The nomograms for breastfed neonates were based on a sample of 108,907 newborns delivered at 14 Kaiser Permanente medical centers in Northern California (United States). The objective of this study was to externally validate the published nomograms for newborn weight loss using data from a geographically distinct population.
Materials and Methods: Data were compiled from the Penn State Milton S. Hershey Medical Center located in Hershey, PA. For singleton neonates delivered at ≥36 weeks of gestation between January 2013 and September 2014, weights were obtained between 6 hours and 48 hours (vaginal delivery) or 60 hours (cesarean delivery) for neonates who were exclusively breastfeeding. Quantile regression methods appropriate for repeated measures were used to estimate 50th, 75th, 90th, and 95th percentiles of weight loss as a function of time after birth. These percentile estimates were compared with the published nomograms.
Results: Of the 1,587 newborns who met inclusion criteria, 1,148 were delivered vaginally, and 439 were delivered via cesarean section. These newborns contributed 1,815 weights for vaginal deliveries (1.6 per newborn) and 893 weights for cesarean deliveries (2.0 per newborn). Percentile estimates from this Penn State sample were similar to the published nomograms. Deviations in percentile estimates for the Penn State sample were similar to deviations observed after fitting the same model separately to each medical center that made up the Kaiser Permanente sample.
Conclusions: The published newborn weight loss nomograms for breastfed neonates were externally validated in a geographically distinct population.

Introduction

Nomograms that show hour-by-hour percentiles of weight loss during the birth hospitalization were recently developed by Flaherman et al.[1] to aid clinical care of well newborns. The nomograms for exclusively breastfed newborns were based on a sample of 108,907 newborns delivered at 14 Northern California (U.S.) Kaiser Permanente medical centers.[1] The publication provided important normative data for evaluating weight loss among breastfed newborns and overcame methodologic issues that plagued previous attempts to quantify newborn weight loss.[2] Nomograms are publicly available at www.newbornweight.org and allow clinicians, lactation support specialists, and others to plot weight loss for an individual newborn and compare against the percentile curves provided by the study. Some have suggested that the nomograms can help alleviate anxieties of new mothers, can help identify newborns with potential breastfeeding difficulties, and may become widely used by clinicians.[3]

An important consideration for wide clinical adoption is whether these nomograms represent patterns of weight loss in other populations. Although the sample population from Kaiser Permanente was diverse with large Hispanic and Asian representation, one limitation is that results may differ for populations with a different race/ethnic mix of newborns.[1] Geographic differences may also be particularly important, especially with respect to breastfeeding. Studies have shown that rates of breastfeeding vary according to geographic regions in the United States[4] and among countries.[5] Factors related to healthcare provider practices such as use of maternal epidural medication,[6] timing and amount of

Departments of [1]Public Health Sciences and [5]Pediatrics, Penn State College of Medicine, Hershey, Pennsylvania.
Departments of [2]Pediatrics and [3]Epidemiology and Biostatistics, School of Medicine, University of California, San Francisco, California.
[4]Division of Research, Kaiser Permanente Northern California, Oakland, California.

maternal intravenous fluids,[7] and intrapartum fluid balance have also been shown to be associated with neonatal weight loss[8] and may vary by hospital or geographic region. Differences among populations with respect to geography, rates of breastfeeding initiation, and other factors conceivably may contribute to different patterns of weight loss.

The Kaiser Permanente data demonstrated consistency between percentile estimates across birth year, when excluding late preterm, postterm, and small- and large-for-gestational-age newborns and after a sensitivity analysis that examined the impact of censoring newborn weights on formula use.[1] However, an external validation to ascertain that the nomograms work well in other populations has not yet been published. We sought to perform this validation by evaluating the weight loss nomograms in a population of newborns at a single medical institution in a geographically distinct region.

Materials and Methods

Participants and outcomes

Data for this study were obtained from electronic medical records at the Penn State Milton S. Hershey Medical Center (Penn State), a U.S. academic medical center located in south central Pennsylvania, in Hershey. For the time period spanning January 2013 to September 2014, data were extracted for all singleton neonates born at ≥36 weeks of gestation with birth weight between 2,000 and 5,000 g (inclusive) who survived to discharge and did not require a neonatal intensive care unit stay during the birth hospitalization. The following data were extracted: type of delivery (vaginal or cesarean section), birth weight, gestational age, length of hospital stay, all weights recorded after birth and during the birth hospitalization, and all instances and timing of formula use during the birth hospitalization. Neonates were excluded if no weights were recorded during the birth hospitalization or if formula was introduced prior to any weight recorded. These inclusion/exclusion criteria matched those from Flaherman et al.[1]

Neonates were typically weighed daily during the birth hospitalization. These daily weights were used to calculate percentage weight changes from birth. Consistent with the technique of Flaherman et al.,[1] neonates were excluded if weight loss was >10% in the first 24 hours or >15% thereafter, or if weight gain was >5% at any time point. These values were considered implausible to occur in the population and were assumed to be data collection or entry errors.

The final sample for analysis included all neonates who met the inclusion and exclusion criteria and who had at least one weight recorded between 6 and 48 hours (vaginal deliveries) or between 6 and 60 hours (cesarean deliveries) while exclusively breastfeeding. A more narrow time window was included in the current study because only 57 and 115 weights were observed >48 hours and >60 hours for vaginal and cesarean deliveries, respectively.

This study was approved by the institutional review board of the Penn State College of Medicine.

Statistical analysis

We applied the same penalized fixed-effects quantile regression model[9] used in Flaherman et al.[1] to estimate the 50th, 75th, 90th, and 95th percentiles of weight loss for the Penn State sample separately for vaginal and cesarean deliveries. The model was specified in exactly the same way as applied to the Kaiser Permanente sample with one exception: only 3 degrees of freedom were used for the B-spline basis[10] that generated nonlinear percentile curves (instead of 4) because the Penn State data included a more narrow time window. The tuning parameter was set to the same value of 5. We graphically compared the estimated percentile curves from the Penn State sample with the Kaiser Permanente sample by plotting both sets of curves on the same figure. We computed the average difference between studies for each percentile curve by calculating the absolute difference between percentile estimates at each integer time point and averaging across these time points.

We hypothesized that observed differences in percentile estimates between studies reflected ordinary and expected random variation. To investigate this hypothesis, we fitted the same quantile regression model separately to each of the 14 medical centers that made up the Kaiser Permanente sample. The panel of fitted percentile curves provided an estimate of sampling variability, or the typical variation that a center might exhibit relative to the overall estimate. We graphically compared the estimated percentile curves from the Penn State sample with this estimate of sampling variability.

Results

For the 2,511 neonates that met inclusion criteria for the Penn State data, 1,587 had at least one weight recorded during the eligible time frame (6–48 hours for vaginal deliveries; 6–60 hours for cesarean deliveries) while exclusively breastfeeding and thus were included in the final sample for analysis. Among the 924 neonates excluded from the analysis, 877 (95%) were excluded because formula was introduced prior to any weight recorded within the eligible time frame, 30 (3%) were excluded because no weight was recorded, and 17 (2%) were excluded because weight loss was considered implausible (e.g., >15% weight loss).

The exclusion with respect to formula use revealed an important difference between the study samples. Among the 2,511 neonates who met inclusion criteria, 706 were excluded due to formula use within 6 hours of birth, a rate of 28.1%. In comparison, the rate of formula use within 6 hours among the Kaiser Permanente sample was 12.6%.[1]

For the 1,587 neonates included in the final sample, 1,148 (72%) were delivered vaginally, and 439 (28%) were delivered by cesarean section. Maternal race/ethnicity was not available in the electronic medical record. However, the primary geographic area served by Penn State had the following race/ethnicity distribution in 2013 according to the U.S. Census Bureau: 68.9% white non-Hispanic, 18.6% black non-Hispanic, 7.9% Hispanic, and 3.6% Asian.[11] These rates were substantially different than the maternal race/ethnicity distribution in the Kaiser Permanente sample (42.7% white non-Hispanic, 6.5% black non-Hispanic, 24.0% Hispanic, and 23.7% Asian).[1] Gestational age, birth weight, and hospital length of stay are summarized in Table 1. The median birth weight for vaginal deliveries among the Penn State sample was 3,400 g, a difference of only 2 g compared with the Kaiser Permanente sample (3,402 g); for cesarean deliveries, the median birth weights were equal (3,470 g).[1] Median gestational age was slightly older for vaginal deliveries at Penn State (40 weeks versus 39 weeks), but cesarean deliveries had

Table 1. Clinical Characteristics by Type of Delivery

	Vaginal (n = *1,148*)	*Cesarean* (n = *439*)
Gestational age (weeks)		
Mean (SD)	39.5 (1.2)	39.2 (1.2)
Median	40	39
Interquartile range	39–40	39–40
Range	36–42	36–42
Birth weight (g)		
Mean (SD)	3,403.0 (432.3)	3,458.8 (481.6)
Median	3,400	3,470
Interquartile range	3,130–3,690	3,140–3,760
Range	2,040–4,960	2,240–4,850
Hospital length of stay (days)		
Mean (SD)	2.0 (0.5)	2.9 (0.6)
Median	2	3
Interquartile range	1.7–2.2	2.6–3.2
Range	1.0–6.4	1.1–5.5

SD, standard deviation.

a median of 39 weeks in both samples.[1] Median hospital lengths of stay were longer at Penn State (2 days for vaginal and 3 days for cesarean section) compared with Kaiser Permanente (1.6 days for vaginal and 2.6 days for cesarean section).[1]

Newborns contributed 1,815 weights for vaginal deliveries (1.6 per newborn) and 893 weights for cesarean deliveries (2.0 per newborn) during the respective eligible time frames. Figure 1 presents the estimated percentile curves of weight loss stratified by delivery type for the Penn State sample compared with the Kaiser Permanente sample.[1] The curves were generally similar. For vaginal deliveries, the Penn State sample had slightly higher estimates at time points near 48 hours for all percentiles and had lower estimates between 12 and 30 hours for the 95th percentile. The single largest difference between studies occurred at 48 hours for the 75th percentile, in which the estimate from the Penn State sample was 7.8% weight loss and the estimate from the Kaiser Permanente sample was 8.3%,[1] a difference of 0.5%. The average differences between the studies ranged from 0.1% (50th percentile) to 0.2% (95th percentile). For cesarean deliveries, higher estimates were observed for the 50th percentile near 60 hours and for the 90th and 95th percentiles near 6 hours, whereas lower estimates were observed for the 90th and 95th percentiles between 12 and 36 hours. The largest difference between studies was also 0.5% for cesarean deliveries; it occurred at 30 hours for the 95th percentile in which the weight loss estimates were 8.5% for the Penn State sample and 8.0% for the Kaiser Permanente sample.[1] The average differences ranged from 0.1% (75th percentile) to 0.2% (95th percentile) for cesarean deliveries.

For each type of delivery, Figures 2 and 3 show the estimated percentile curves in the Penn State sample compared with the sampling variability among the 14 Kaiser Permanente medical centers.[1] Sample sizes for these medical centers ranged from 2,140 to 12,060 for vaginal deliveries and 846 to 3,083 for cesarean deliveries,[1] much larger than the Penn State sample. Figures 2 and 3 show that differences in percentile estimates from the Penn State sample were comparable to differences that occurred among the Kaiser Permanente medical centers.

Discussion

Results from the current study demonstrated that the newborn weight loss nomograms for exclusively breastfed newborns published by Flaherman et al.[1] generalized to a

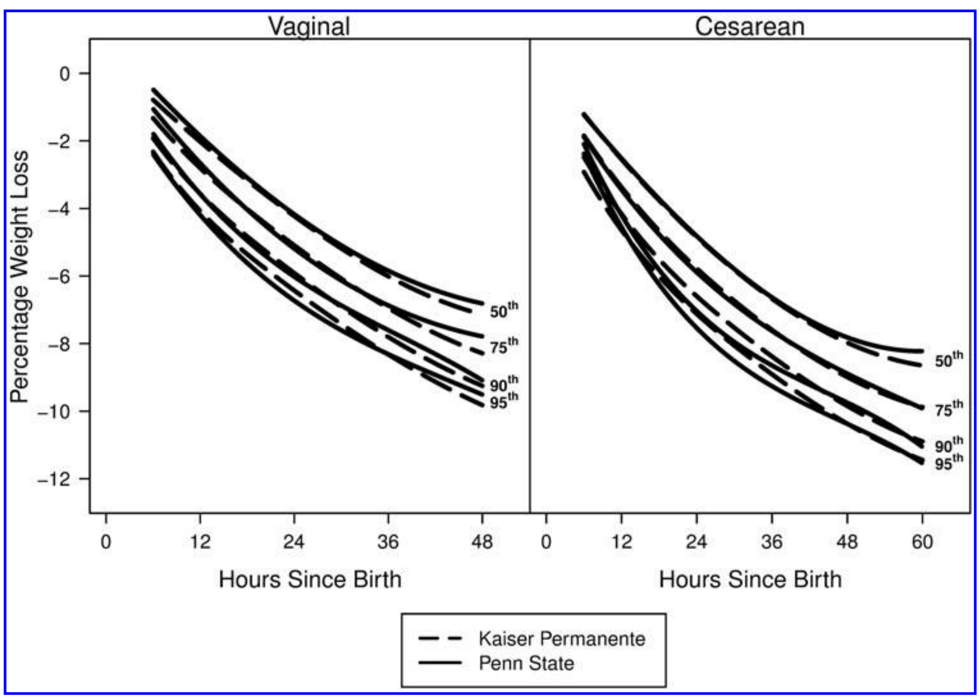

FIG. 1. Estimated percentile curves of percentage weight loss for the Penn State sample (solid curve) compared with the Kaiser Permanente sample[1] (dashed curve) stratified by type of delivery.

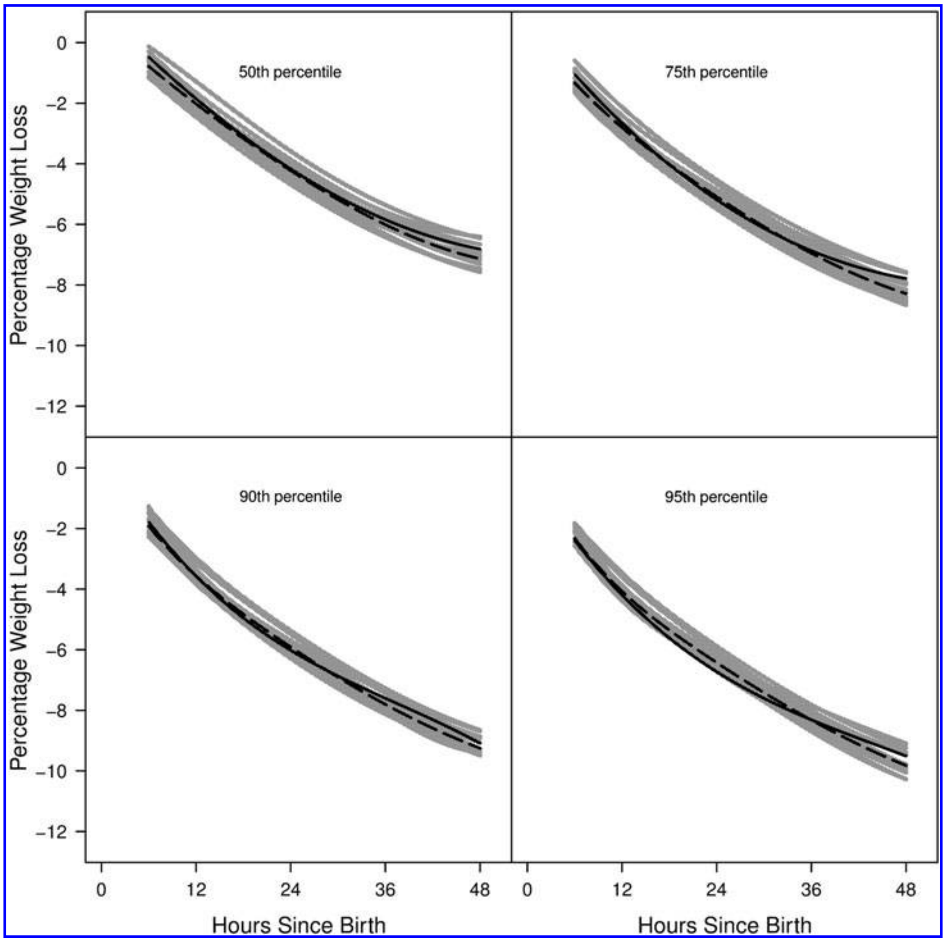

FIG. 2. Comparison of percentile estimates from the Penn State sample (solid black line) with the Kaiser Permanente sample (dashed black line) and each of the 14 centers that made up the Kaiser Permanente sample[1] (gray lines) for newborns delivered vaginally.

geographically distinct population. Similarities between populations were evident when plotting percentile estimates on the same figure. Average and maximum differences in percentile estimates between populations were small for each type of delivery and were consistent with sampling variability among the medical centers that made up the Kaiser Permanente sample. Although no single statistical test exists to test whether a model is externally validated, a validation study such as this provides an unbiased estimate of model performance that is used to make informed clinical judgments.[12] Our conclusion is that the differences observed in this study were clinically insignificant, and thus the newborn weight loss nomograms generated from data from Kaiser Permanente medical centers were externally validated.

The primary implication of this study is that the published newborn weight loss nomograms can be used with greater confidence in different geographic areas, in populations with different race/ethnicity compositions, and in regions with different rates of formula supplementation. Furthermore, this study suggests that differences in health system–related clinical practices are unlikely to affect the utility of the nomograms. Study populations also likely differed in numerous other ways not captured in these datasets, such as maternal parity, marital status, prenatal care, obstetric care, or amount of lactation support during the birth hospitalization. Study populations were similar with respect to gestational age and birth weight, with median estimates largely in agreement with U.S. national estimates for newborns whose mothers intended to breastfeed, although median birth weights were slightly larger.[13] The similarity of both study populations to national estimates supports a supposition that the newborn weight loss nomograms published by Flaherman et al.[1] are accurate for newborn populations in the United States.

Two studies were recently published that also estimated percentiles of weight loss as a function of time after birth, although neither was meant to be a validation study. A study by Fonseca et al.[14] included 1,288 newborns from five maternity units in Portugal, and a study by Bertini et al.[15] included 1,760 newborns at a single Italian medical center that ‘‘focuses on natural childbirth, without medical assistance.’’ The percentile estimates in Fonseca et al.[14] were obtained for all newborns regardless of delivery type and differed only slightly from the much larger Kaiser Permanente sample described by Flaherman et al.[1]; estimates of weight loss for the 90th percentile for 6, 12, 24, 36, and 48 hours after birth were an average of 1.2% and 0.6% lower[14] than the respective estimates for vaginal and cesarean deliveries for the Kaiser Permanente sample.[1] In contrast, the results differed

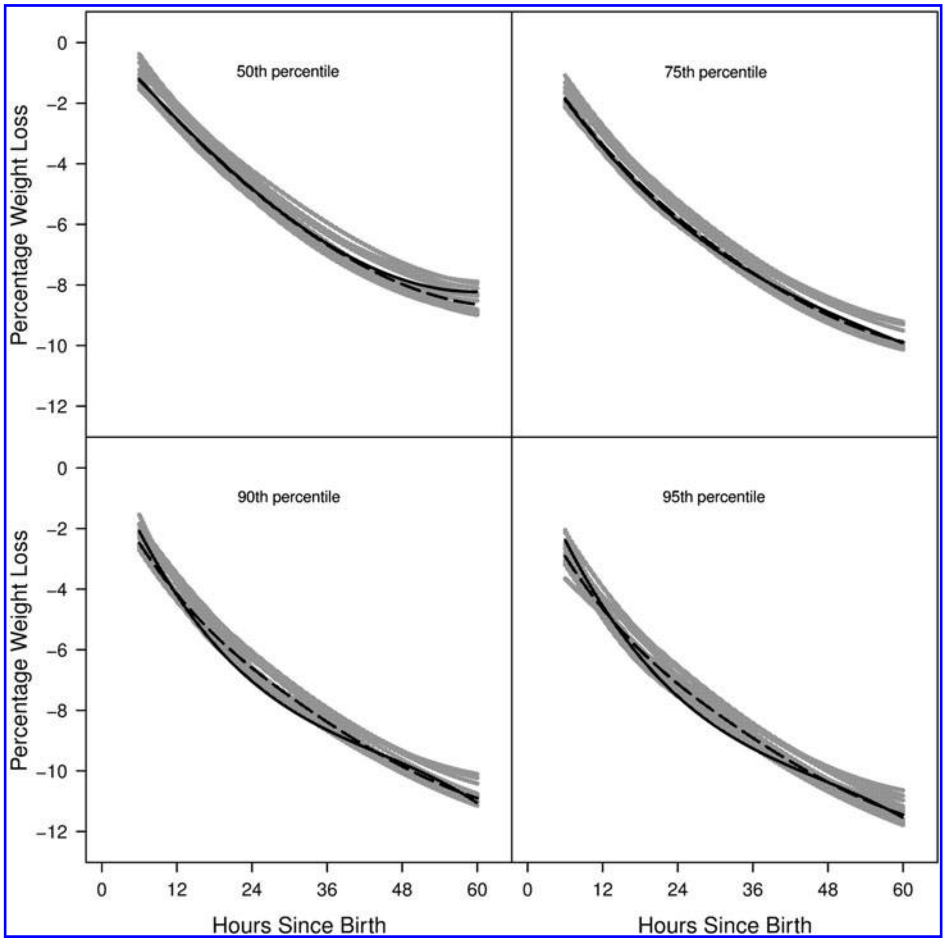

FIG. 3. Comparison of percentile estimates from the Penn State sample (solid black line) with the Kaiser Permanente sample (dashed black line) and each of the 14 centers that made up the Kaiser Permanente sample[1] (gray lines) for newborns delivered via cesarean section.

substantially for Bertini et al.[15] For example, the percentage weight loss at 48 hours for the 95th percentile was 8.0% in Bertini et al.[15] versus 9.8% from the Kaiser Permanente sample[1] for vaginal deliveries.

Reasons for these differences were unlikely to be related to population characteristics such as gestational age and birth weight because these variables were generally similar to the Kaiser Permanente sample in each study.[1,14,15] The lower estimates in Fonseca et al.[14] may be due to including weights indicating >15% weight loss, which Flaherman et al.[1] considered to be likely erroneous values, as well as the statistical methods used by Fonseca et al.,[14] which required an important assumption of conditional normality and imposed a polynomial structure on all percentile curves,[16,17] neither of which needed to be assumed for the statistical methods used in Flaherman et al.[1] The statistical methods in Bertini et al.[15] were related to the methods used in Flaherman et al.[1] and are unlikely to explain the differences between the studies.

Rather, important differences between standard practices at the medical centers included in each study may explain the differences. For example, Bertini et al.[15] described a focus on natural childbirth practices that may include less maternal intrapartum fluid administration, which has been shown to be associated with lower newborn weight loss.[7,8]

Our study results are limited in three ways. First, the analysis used a narrower time window than the published nomograms because our sample size was not large enough to include enough weights at later time points. Although the time window included in the study represented 97% of weights for vaginal deliveries and 91% of weights for cesarean deliveries observed in the Kaiser Permanente sample,[1] nevertheless no claims from this study can be made regarding validation after 48 or 60 hours. Second, the study provides no guarantee that the newborn weight loss nomograms will be accurate in all other clinical settings. However, this is a general limitation of all validation studies.[18] Third, the same study team conducted both the initial and validation studies. Some authors have suggested that validation studies conducted by the same study team are more likely to be reported with inappropriately favorable interpretations.[19]

Conclusions

Newborn weight loss nomograms published by Flaherman et al.[1] meant to aid clinical care for exclusively breastfed newborns were externally validated in this study. The study provides additional confidence that these nomograms accurately reflect weight loss for exclusively breastfed newborns.

Acknowledgments

This work was supported by grant R40MC26811 from the Maternal and Child Health Bureau, Health Resources and Services Administration, U.S. Department of Health and Human Services.

Disclosure Statement

I.M.P. reports personal fees from Procter & Gamble Company, Pfizer, Perrigo Nutritionals, and McNeil Consumer Healthcare, outside the submitted work. E.S., V.J.F., M.W.K., S.X.L., and E.M.W. declare no competing financial interests exist.

References

1. Flaherman VJ, Schaefer EW, Kuzniewicz MW, et al. Early weight loss nomograms for exclusively breastfed newborns. *Pediatrics* 2015;135:e16–e23.
2. Noel-Weiss J, Courant G, Woodend AK. Physiological weight loss in the breastfed neonate: A systematic review. *Open Med* 2008;2:e99–e110.
3. Taylor JA, Simpson EA. Categorizing weight loss in breastfed infants: A good first step. *Pediatrics* 2015;135:e174–e175.
4. Kogan MD, Singh GK, Dee DL, et al. Multivariate analysis of state variation in breastfeeding rates in the United States. *Am J Public Health* 2008;98:1872–1880.
5. Cai X, Wardlaw T, Brown DW. Global trends in exclusive breastfeeding. *Int Breastfeed J* 2012;7:12.
6. Martens PJ, Romphf L. Factors associated with newborn in-hospital weight loss: Comparisons by feeding method, demographics, and birthing procedures. *J Hum Lact* 2007;23:233–241, quiz 242–235.
7. Noel-Weiss J, Woodend AK, Peterson WE, et al. An observational study of associations among maternal fluids during parturition, neonatal output, and breastfed newborn weight loss. *Int Breastfeed J* 2011;6:9.
8. Chantry CJ, Nommsen-Rivers LA, Peerson JM, et al. Excess weight loss in first-born breastfed newborns relates to maternal intrapartum fluid balance. *Pediatrics* 2011;127:e171–e179.
9. Koenker R. Quantile regression for longitudinal data. *J Multivariate Anal* 2004;91:74–89.
10. Koenker R, Ng PIN, Portnoy S. Quantile smoothing splines. *Biometrika* 1994;81:673–680.
11. U.S. Census Bureau. Dauphin County, PA. Available at http://quickfacts.census.gov/qfd/states/42/42043.html (accessed May 8, 2015).
12. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000;19:453–473.
13. National Center for Health Statistics. Vital Statistics Natality Birth Data. 2013. Available at www.nber.org/data/vital-statistics-natality-data.html (accessed May 6, 2015).
14. Fonseca M, Severo M, Santos A. A new approach to estimating weight change and its reference intervals during the first 96 hours of life. *Acta Paediatr* 2015;104:1028–1034.
15. Bertini G, Breschi R, Dani C. Physiological weight loss chart helps to identify high-risk infants who need breastfeeding support. *Acta Paediatr* 2015;104:1024–1027.
16. Royston P. Calculation of unconditional and conditional reference intervals for foetal size and growth from longitudinal measurements. *Stat Med* 1995;14:1417–1436.
17. Altman DG. Construction of age-related reference centiles using absolute residuals. *Stat Med* 1993;12:917–924.
18. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;130:515–524.
19. Collins GS, de Groot JA, Dutton S, et al. External validation of multivariable prediction models: A systematic review of methodological conduct and reporting. *BMC Med Res Methodol* 2014;14:40.

Address correspondence to:
Eric W. Schaefer, MS
Department of Public Health Sciences
Penn State College of Medicine
500 University Drive, A210
Hershey, PA 17033

E-mail: ews12@psu.edu