



Is routine evaluation of gastric residuals for premature infants safe or effective?

Matthew A. Rysavy¹ · Patricia L. Watkins² · Tarah T. Colaizy¹ · Abhik Das³

Received: 3 December 2019 / Revised: 11 December 2019 / Accepted: 19 December 2019 / Published online: 7 January 2020
© Springer Nature America, Inc. 2020

Manuscript citation: Parker LA, Weaver M, Murgas Torrazza RJ, Shuster J, Li N, Krueger C, Neu J. Effect of gastric residual evaluation on enteral intake in extremely preterm infants: a randomized clinical trial. *JAMA Pediatr.* 2019;173:534-43.

Infants with congenital/chromosomal anomalies were excluded.

Intervention: Routine evaluation of prefeed gastric residuals versus no evaluation of prefeed gastric residuals. Gastric content aspiration was not used to verify feeding tube placement for infants in the no evaluation group.

Type of investigation

Treatment.

Question

In infants born at ≤ 32 weeks' gestational age and ≤ 1250 grams, does omitting routine prefeed gastric residual evaluation compared with routine prefeed gastric residual evaluation affect enteral nutrition intake during the first 6 weeks after birth?

Methods

Design: Randomized controlled trial, unblinded.

Setting: Level IV neonatal intensive care unit (NICU).

Patients: Infants born at ≤ 32 weeks' gestational age and ≤ 1250 g, < 72 h old, and receiving feedings for < 24 h.

Outcomes

Primary outcome: Week-average enteral nutritional intake (ml/kg/d) for the first 6 weeks of life. This is presented as co-primary outcomes of week-averaged daily feeding volumes (six comparisons) and the average difference in weekly feeding advancement between no gastric residual evaluation and evaluation groups (assessed using a treatment \times postnatal week interaction on week-average enteral intake).

Secondary outcomes: Weekly weight, head circumference, and length; days to full enteral nutrition (120 ml/kg/d); hours of parenteral nutrition; hours of central line exposure; levels of serum direct bilirubin and alkaline phosphatase; length of stay; episodes of emesis; episodes of abdominal distension; number of abdominal radiographs; episodes of presumed or culture-proven sepsis; pepsin level in weekly tracheal aspirate; necrotizing enterocolitis (NEC); and death. The trial protocol did not prospectively define several secondary outcomes that were presented in the manuscript (bronchopulmonary dysplasia, intraventricular hemorrhage, days of invasive mechanical ventilation and use of any respiratory support at 6 weeks). The protocol did define some secondary outcomes that were not presented in the manuscript (positive stool guaiac, fecal calprotectin, and serum levels of gastrin and motilin).

Most outcomes were obtained through 6 weeks' postnatal age (e.g., death and NEC), although a few were obtained through hospital discharge (e.g., length of stay). Overall, 28 of 146 (19%) infants were not followed through

✉ Matthew A. Rysavy
matthew-rysavy@uiowa.edu

¹ Division of Neonatology, Stead Family Department of Pediatrics, University of Iowa, Iowa City, IA, USA

² Division of Neonatology, Department of Pediatrics, Sanford Medical Center, Bismarck, ND, USA

³ Biostatistics and Epidemiology Division, RTI International, Rockville, MD, USA

6 weeks due to elective study withdrawal, death, NEC, or spontaneous intestinal perforation.

Analysis and sample size: The primary outcome was analyzed using generalized linear mixed modeling with week-average daily enteral nutritional intake (ml/kg/d) as a continuous outcome for each week. In unadjusted analyses, for 80% power to detect a 50% difference from an estimated 72 ml/kg/d (i.e., to 108 ml/kg/d) during a given week, 104 infants were required for a two-sided t test to detect $p < 0.05$ without correction for one interim analysis or multiple outcome comparisons.

Main results

After post hoc adjustment for birthweight, gestational age at birth, and a gestational age at birth \times postnatal week interaction, the authors showed a difference in advancement of 20.7 versus 17.9 ml/kg/d between the no gastric residual evaluation and evaluation groups ($p = 0.02$). There were statistically significant differences in unadjusted average daily enteral nutrition at weeks 5 and 6 (137.2 vs 123.9 mL/kg/d, and 141.6 vs 128.4 ml/kg/d, respectively), but not at earlier weeks.

Secondary outcomes were analyzed using several multivariable modeling strategies. Variables included in the models were determined post hoc and selected based on observed relationships in the data. As an example: After adjustment for birthweight, gestational age at birth, postnatal week, and a gestational age at birth \times treatment group interaction term and log-transformation of the birthweight variable, average weight during the 6 weeks was greater in the no gastric residual evaluation group (1145.3 vs 1129.7 g, $p = 0.03$). This result was statistically significant but of unclear clinical significance.

Study conclusion

Enteral feeding advanced more rapidly in the group of infants randomized to no routine gastric residual evaluation. This group also demonstrated higher feeding volumes at 5 and 6 weeks' postnatal age. The authors did not detail the management of gastric residuals (e.g., discarded versus refeed), which may have affected the primary outcome. The study was not powered for important outcomes such as mortality or NEC.

Commentary

NICUs worldwide commonly incorporate routine evaluation of prefeed gastric residuals into their management of preterm infants. Preterm infants may undergo this procedure

eight or more times daily for weeks or months after birth. Large or discolored residuals are thought to suggest "feeding intolerance" and have been associated, in case-control studies, with NEC [1, 2]. Theoretically, large gastric residuals may also lead to pulmonary aspiration and ventilator-associated pneumonia [3].

The effectiveness and safety of routine checks of prefeed gastric residuals have not been established. Perceptions of what constitutes a "large" residual vary widely, and there are no standards to guide how residuals should be handled after aspiration (e.g., discarded, refeed without effect on the subsequent feed, or refeed and subtracted from the subsequent feed). Gastric emptying may be affected by infant maturity, medications (e.g., methylxanthines, mydriatics), feed composition (formula versus breastmilk), feeding method (bolus versus continuous feedings), infant position, tube position, and infant state (e.g., respiratory distress, infection). Theoretically, evaluation of gastric residuals may cause harm: for example, repetitive negative pressure from aspiration may injure sensitive gastric mucosa, and discarding gastric residuals may result in the loss of essential gastric enzymes, gastric acid, and important nutrients [4]. Misperception of the significance of gastric residuals may lead to unnecessary restriction of feeding volumes or unnecessary blood draws or radiographs.

Two other clinical trials have compared the routine evaluation of prefeed gastric residuals to forgoing the practice: In a smaller trial of 61 similar infants performed at the same NICU during the 3 years prior to the current trial, the authors found no statistically significant differences in amount of feeding at 2 or 3 weeks' age, growth, or parenteral nutrition [5]. In a small trial of 88 larger infants (1500–2000 g), a separate group of authors found no statistically significant differences in time to reach full feeds (120 ml/kg/d), time to regain birthweight, episodes of feeding interruptions or sepsis between routine evaluation and no evaluation groups [6].

The present trial had additional power to detect differences in outcomes between the trial arms because it was both larger ($n = 143$) than previous trials and adjusted for prognostic variables in primary and secondary analyses. The authors found that infants randomized to no routine gastric residual evaluation advanced enteral feeding more rapidly than those randomized to no residual evaluation.

An important limitation of clinical trials that randomize a treatment that alters the availability of clinical information (here, prefeed gastric residual volume and color) between groups is that the clinical team must be unblinded to the intervention. This may have led to conscious or unconscious biases in clinical decisions (e.g., daily feeding advancement) that affected study outcomes (e.g., daily enteral feeding volumes). Among infants who underwent routine evaluation for gastric residuals, the authors did not

specify how residuals were handled. Differences in approach (discarding versus refeeding) could systematically and independently affect enteral nutritional intake in the residual group. It is noteworthy that 18 of 69 (26%) infants randomized to no routine gastric residual evaluation had ≥ 1 residuals evaluated for signs of gastrointestinal dysfunction, leading to crossover between the study groups. Also, the trial did not enroll an adequate number of infants to detect potentially meaningful differences in important outcomes such as death or NEC.

This trial highlights important areas of uncertainty about a common practice for which there is insufficient evidence for efficacy or safety and establishes foundations for future studies to build upon. However, the practice of obtaining prefeed gastric residuals deserves further evaluation before NICUs abandon the practice entirely. Until other studies are reported, we cannot confidently answer major questions surrounding evaluation of gastric residuals for preterm infants (Among whom? Under which circumstances? How to act upon the information obtained?).

EBM lesson: multivariable models in the analysis of clinical trials

Well-conducted randomized clinical trials, when ethically and logistically feasible, provide the optimal study design to measure true causal effects of a medical intervention. Although clinicians may be generally aware of how decisions about eligibility, allocation, and outcome assessment affect trial interpretation, the impact of choices about statistical analyses may be less appreciated. Multivariable analyses—that is, when the effect of a primary variable (e.g., treatment allocation) is adjusted to account for the effect of other variables—may be familiar to clinicians from observational studies where patients are not randomly and prospectively assigned to an intervention. However, such analyses are also frequently used in randomized clinical trials.

Reasons to perform multivariable analyses in the analysis of trial outcomes include: [7]

Obtaining appropriate confidence intervals and p values after stratified randomization

Some trials use stratification methods to ensure that equal numbers of infants at a given center (in a multicenter trial) or from certain birthweight or gestational age groups are randomized to each arm (e.g., when stratified by center, performed so that a center that enrolls $\sim x$ infants in the intervention arm also enrolls $\sim x$ infants in the control arm). The effect of the variables used for stratification should be accounted for by using a multivariable analysis of the primary outcome to avoid loss of

study power and inappropriately large p values and wide confidence intervals for the treatment effect [8].

Accounting for chance imbalance in important baseline variables

When stratified randomization is impractical, consideration is sometimes given to balancing important prognostic variables in the analysis phase only. While randomization promotes on-average balance of measured and unmeasured variables across treatment groups, it does not ensure balance of any particular variable. Importantly, the trial protocol should prospectively define multivariable adjustment planned for the primary analysis. *Post hoc* selection of variables may increase the risk of false-positive findings (type I error) [9, 10].

Increased power to detect differences in outcomes

Adjustment for prognostic variables (i.e., those that affect the outcome) in clinical trial analyses eliminates some of the variation in outcomes among patients unrelated to treatment assignment. For continuous outcome variables, multivariable adjustment reduces the standard error in the treatment effect (i.e., potentially makes the treatment groups less alike). In analyses of dichotomous (e.g., mortality) or time-to-event (e.g., days until full feeds) outcomes, multivariable adjustment results in an increase in standard error (which may lead to wider confidence intervals), but this is usually compensated for by an increase (away from the null) in the measured effect size. In both circumstances, trial power—the ability to detect a statistically significant difference in outcomes between groups when one exists—is increased [7, 11].

The CONSORT statement, which specifies practices for the conduct and reporting of clinical trials required by many medical journals, states that adjusted analysis of clinical trials “may be sensible, especially if one or more variables is thought to be prognostic,” recommends pre-specifying adjusted analyses in the study protocol, and encourages multivariable adjustment for variables used for stratification in trial design [12].

Acknowledgements The Journal Club is a collaboration between the American Academy of Pediatrics - Section of Neonatal Perinatal Medicine and the International Society for Evidence-Based Neonatology (EBNEO.org).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Cobb BA, Carlo WA, Ambalavanan N. Gastric residuals and their relationship to necrotizing enterocolitis in very low birth weight infants. *Pediatrics*. 2004;113:50–3.
2. Bertino E, Giuliani F, Prandi G, Coscia A, Martano C, Fabris C. Necrotizing enterocolitis: risk factor analysis and role of gastric residuals in very low birth weight infants. *J Pediatr Gastroenterol Nutr*. 2009;48:437–42.
3. Farhath S, Aghai ZH, Nakhla T, Saslow J, He Z, Soundar S, et al. Pepsin, a reliable marker of gastric aspiration, is frequently detected in tracheal aspirates from premature ventilated neonates: relationship with feeding and methylxanthine therapy. *J Pediatr Gastroenterol Nutr*. 2006;43:336–41.
4. Li YF, Lin HC, Torrazza RM, Parker L, Talaga E, Neu J. Gastric residual evaluation in preterm neonates: a useful monitoring technique or a hinderance? *Pediatr Neonatol*. 2014;55:335–40.
5. Torrazza RM, Parker LA, Li Y, Talaga E, Shuster J, Neu J. The value of routine evaluation of gastric residuals in very low birth weight infants. *J Perinatol*. 2015;35:57–60.
6. Singh B, Rochow N, Chessell L, Wilson J, Cunningham K, Fusch C, et al. Gastric residual volume in feeding advancement in preterm infants (GRIP Study): a randomized trial. *J Pediatr*. 2018;200:79–83.
7. Kahan BC, Jairath V, Doré CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. *Trials*. 2014;15:139.
8. Kahan BC, Morris TB. Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis. *BMJ*. 2012;345:e5840.
9. Altman DG (2005). Covariate imbalance, adjustment for. In: Armitage P, Colton T, editors. *Encyclopedia of Biostatistics*. 2nd ed. Chichester:John Wiley & Sons, Ltd; 2005, p. 1273–8.
10. Raab GM, Day S, Sales J. How to select covariates to include in the analysis of a clinical trial. *Control Clin Trials*. 2000;21:330–42.
11. Hernández AV, Steyerberg EW, Habbema DF. Covariate adjustment in randomized controlled trials with dichotomous outcomes increases statistical power and reduces sample size requirements. *J Clin Epidemiol*. 2004;57:454–60.
12. Moher D, Hopewell S, Shulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.