We thank Dr. Paneth for his insightful comments (1) and, in particular, the way in which he has located our work on fetal growth (2) in the context of the larger body of work on fetuses at risk. We would like to respond to some of the concerns raised regarding the measurement of intrauterine growth, as well as touch briefly on the larger issue of selecting denominators in perinatal epidemiology.

We agree entirely with Dr. Paneth’s criticism (1) of the focus on dichotomous measures of fetal growth (“small-for-gestational-age” vs. “appropriate-for-gestational-age”) instead of continuous ones such as birth weight \( z \) scores. However, as with percentiles, we argue that it is important to ensure that the mean weights (and standard deviations) used to calculate \( z \) scores are based on the average weights of all fetuses that progressed to a given gestational age, not the average weights of only fetuses subsequently born that week. We agree that epidemiologists would be well served to respect the natural continuum of fetal growth in their attempts to better understand the etiology of growth restriction, but they should do so with a measure that is not associated with gestational age at birth.

The limitations of existing intrauterine weight standards, including small sample sizes and errors in estimation of fetal weight, are important to recognize. We should clarify that the goal of our calculations was not to establish the precise extent of the bias arising from the use of conventional birth weight percentiles but rather to help understand whether the magnitude of the theoretical bias we outlined was of substantive importance. We had also repeated our simulations using a Canadian birth weight reference (3) and ultrasound data from the Royal Victoria Hospital population presented in figure 1 (2) (a sample of 3,015 routine 32-week ultrasounds once female fetuses were included, which, of note, is a larger sample than the 1,792 births at 32 weeks used to create the Canadian reference). This method produced a similar, if slightly larger, bias than that obtained from our simulations based on Norwegian data. This finding, combined with the observation that the 32-week estimated fetal weights of fetuses later born preterm in our population were already significantly lower than the estimated fetal weights of fetuses subsequently born at term (confirming previous reports from a smaller study (4)), led us to believe that the bias presented in our paper was unlikely to be solely an artifact from errors in estimation of fetal weight or fetal weight standards.

The exact magnitude of the bias requires further refinement, however, and research to improve estimated fetal weights formulae and create a new fetal weight standard (5) will be a valuable contribution toward this end. Research is also needed to develop a standard that combines the weights of births and ongoing pregnancies (i.e., all fetuses at risk at the beginning of a given week) so that it is methodologically appropriate for all gestational ages.

Since fetal growth (and growth restriction) occurs in utero, we have argued (2) that all fetuses at risk of being small at a given gestational age should be included in the creation of normative weight ranges. Application of this fetuses-at-risk principle to other areas of perinatal epidemiology is more controversial, and we can appreciate arguments against the use of fetuses at risk for outcomes such as cerebral palsy or infant mortality (1). We would propose, however, that much of the debate on fetuses-at-risk versus total-birth reference groups has suffered from attempts to make generic recommendations to cover all research questions. Perhaps it is time to begin considering each situation individually and to establish the most
appropriate denominators based on the research question, the outcome of interest, and the hypothesized timing and mechanism of the exposure. As for neonatologists, although their clinical focus is on the care of their liveborn charges and not fetuses in utero, we would still propose that the most appropriate standard to use is that which is best able to identify infants at increased risk of adverse neonatal outcomes, irrespective of whether this is obtained through an intrauterine or birth weight standard. Some work has been done to compare the predictive ability of intrauterine and neonatal weight standards for risk of adverse outcomes such as respiratory morbidity (6, 7), but studies with sufficient statistical precision to conclusively distinguish between the two standards are needed before recommendations for clinical use can be made.

ACKNOWLEDGMENTS
Conflict of interest: none declared.

REFERENCES