Appendix 2. Statistical Appendix

Sample size calculation

To determine the sample size, we considered the precision of the data in estimating specific centiles. We assumed data were normally distributed at each gestational age. The SE of the p^{th} centile is given using the standard formula (1):

$$SEp = SD\sqrt{(1+\frac{1}{2}Z_p^2)/n}$$

where SE is the standard error, SD is the standard deviation of the measurement of interest (which may change according to gestational age), Z_p is the value of the standard normal distribution corresponding to the p^{th} centile, and n is the sample size. To create an evidence-based early warning score we desired a 95% CI with an SE of < 0.10*SD at the boundaries.

Our sample size calculations are based on conservative estimates as they do not take account of the effect of serial measurements from the same women (2). Using Royston's design factor of 2.3, a longitudinal study of 1000 women could have equivalent precision to a cross-sectional study of around 2300 women (3).

Approach to obtaining best fit for the data

We explored different statistical methods to achieve the best fit to the data. In brief, these methods included: mean and SD using fractional polynomials (4); Cole's lambda, mu and sigma method (5); the lambda, mu, sigma and Box-Cox t distribution method (6); the lambda, mu, sigma and Box-Cox power exponential method (7); the skew power exponential method (7); the skew t distribution

method; and multilevel models. To present curves, we tried different smoothing techniques including fractional polynomial (7), cubic splines (5) and penalised splines (8).

We assessed goodness of fit by visual inspection of empirical centiles versus fitted centiles, quantile-quantile plots of the residuals, plots of residuals versus fitted values, and the distribution of fitted Z scores across gestational ages.

We used bootstrap resampling methods to estimate 95% confidence intervals for the centiles at two-weekly intervals, with fifty bootstrap replicates. We repeated the analysis on each of the replicates and used the standard deviation of these results as an approximation for the standard error.

References

1. Healy MJR. Notes on the statistics of growth standards. Ann Hum Biol. 1974;1:41–6.
2. Altman DG, Ohuma EO. Statistical considerations for the development of prescriptive fetal and newborn growth standards in the INTERGROWTH-21stProject. BJOG An Int J Obstet Gynaecol. 2013;120(SUPPL. 2):71–6.
3. Royston P. Calculation of unconditional and conditional reference intervals for foetal size and growth from longitudinal measurements. Stat Med. 1995;14:1417–36.
4. Royston P, Altman DG. Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. Appl Stat. 1994;43:429–67.
5. Cole TJ, Green PJ. Smoothing reference centile curves: The lms method and penalized likelihood. Stat Med. 1992;11:1305–19.
6. Rigby RA, Stasinopoulos DM. Using the Box-Cox t distribution in GAMLSS to model skewness and kurtosis. Stat Modelling. 2006;6:209–29.
7. Rigby RA, Stasinopoulos DM. Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. Stat Med. 2004;23:3053–76.
8. Eilers P, Marx B. Flexible smoothing with B -splines and penalties. Stat Sci. 1996;11:89–121.

Green LJ, Mackillop LH, Salvi D, Pullon R, Loerup L, Mossop J et al. Gestation-specific vital sign reference ranges in pregnancy. Obstet Gynecol 2020;135.
The authors provided this information as a supplement to their article.