Assessing the accuracy of ultrasound estimation of gestational age during routine antenatal care in in vitro fertilization (IVF) pregnancies

Bridget Knight¹, Aaron Brereton², Roy J Powell^{3,4} and Helen Liversedge²

Abstract

Introduction: In the UK an accurate gestational age is confirmed by ultrasound measured foetal crown rump length (CRL) at 11 + 2–14 + 1 weeks of gestation. The currently recommended Robinson and Fleming crown rump length reference chart was developed in 1975. Advances in ultrasound technology and standardized crown rump length measurement training could mean this is now out of date. Our study aimed to assess its accuracy in current routine antenatal care.

Methods: Retrospective data from 178 IVF pregnancies seen for routine antenatal care at a UK Regional Maternity Unit between 1 January 2006 and 1 January 2016 was retrieved. We compared ultrasound calculated crown rump length gestational age taken at the routine First Trimester Screening Clinic (FTSC) with the 'true' gestational age calculated from the known IVF fertilization date.

Results: We identified a systematic overestimation of gestational age by ultrasound using the currently recommended crown rump length reference chart when compared to IVF gestational age. The mean overestimation was 3.0 days (95% CI: 2.7 to 3.4), $p < 0.001$. A range of alternative ultrasound reference charts also generated a systematic overestimation, ranging from 1.6 to 2.9 days ($p < 0.001$, for each).

Conclusions: The current crown rump length reference chart systematically overestimates gestational age by an average of three days when assessed in IVF pregnancies. A systematic overestimation was also identified in alternative crown rump length reference charts. These differences, although slight, were systematic with implications for the accuracy of gestational age estimation particularly in pregnancies at risk of pre-term delivery or growth restriction. Our findings need confirming in larger, non IVF cohorts and could lead to the need for an updated crown rump length reference chart.

Keywords

Scanning, pregnancy, research

Date received: 18 September 2017; accepted: 2 December 2017

Introduction

Accurate estimation of gestational age (GA) in early pregnancy is a key component of routine antenatal care. It is important for the mother, who wants to know her expected date of delivery (EDD) and for clinicians to ensure optimal timing for screening tests,

¹NIHR Exeter Clinical Research Facility, University of Exeter Medical School and RD&E NHS FT, Exeter, UK

²Antenatal Ultrasound Department, Centre for Women's Health, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK

³Research Design Service (RDS) South West, Exeter, UK

⁴R&D Department, RD&E NHS FT, Exeter, UK

Corresponding author:

Helen Liversedge, Associate Specialist, Obstetric Ultrasound Department, Centre for Women's Health, Royal Devon & Exeter NHS Foundation Trust, Barrack Rd, Exeter, EX2 5DW, UK.
Email: helen.liversedge@nhs.net

foetal well-being assessments and induction of labour if required.^{1–5} The initial EDD is usually estimated from the mother's menstrual history, where this is known and accurate, followed up by an ultrasound scan (USS). For all mothers in the UK the routine time for this scan is at 11+2–14+1 weeks of gestation: the crown rump length (CRL) of the foetus is measured and compared to the standard reference dating chart to provide an accurate EDD and confirm GA at the time of the scan. A nuchal translucency measurement is also taken at this time if required for the Standard NHS Combined Screening Test.⁶

There are several CRL charts available: a systematic review⁷ of 29 studies reporting CRL estimations of GA identified four reference ranges^{8–11} whose study design met the review authors own derived quality criteria, and the recently proposed INTERGROWTH reference range.¹² The Robinson and Fleming CRL chart⁸ is currently recommended for use by the British Medical Ultrasound Society (BMUS)¹³ and UK National Screening Committee (UKNSC)/Fetal Anomaly Screening Programme (FASP)⁶ and is used both locally and nationally. This reference range was derived in 1975 from menstrual history data given on the day of the scan: menstrual history data is self-reported and often inaccurate. Advances in ultrasound technology have resulted in greatly improved resolution of scanned images; this combined with new standardized training for measuring the CRL mean that the 1975 charts may now be out of date. Our study aimed to assess the accuracy of the current standard reference range used locally to estimate GA in routine antenatal care. We compared GA data obtained routinely at 11+2–14+1 weeks in IVF pregnancies with the 'true' age based on the known date of fertilization.

Methodology

Routinely collected retrospective data from all IVF pregnancies seen for antenatal care at the Royal Devon & Exeter NHS Foundation Trust (RD&E NHS FT) over the period 1 January 2006 to 1 January 2016 were identified from the Clinical Viewpoint database.¹⁴ Inclusion criteria were: singleton IVF pregnancy, fresh 0+embryo cycle, known dates of oocyte retrieval/fertilization, no structural/chromosomal abnormalities and progression to viability (24 weeks gestation). Records with incomplete or missing data were also excluded from analysis. CRL images were reviewed to ensure adequate accuracy. The data for those included were anonymized and downloaded to a study specific database. These included GA (in days) calculated from the date of fertilization, and GA (in days) from the CRL reference chart at the time of 11+2–14+1

week First Trimester Screening Clinic (FTSC) USS. Differences in days between GA calculations (fertilization date vs. USS CRL) were explored with the current UK CRL reference range used in clinical practice,⁸ and then expanded to include other currently recommended reference ranges available.^{9–12}

FTSC ultrasound examination

USS examinations were performed by experienced sonographers or foetal medicine doctors. All were Fetal Medicine Foundation (FMF) and Down's Syndrome Screening Quality Assurance Support Service (DQASS) accredited to perform FTSC screening and are audited six monthly to ensure standardization and accuracy.⁶ Standardized transabdominal (TA) USS's were performed using modern high-resolution ultrasound machines (Toshiba Aplio 500 and XG, GE Voluson E8 or Siemens S2000) as per UKNSC/FASP guidelines.⁶

Sample size calculation

We aimed to obtain a sample size close to 200. A sample size of 100 would provide a 95% CI around the limits of agreement on a Bland–Altman plot of ± 0.34 standard deviations of the differences between measurements. A sample of 200 would provide ± 0.24 standard deviations.¹⁵

Statistical methods

Bland–Altman¹⁵ plots were used initially to identify any systematic/random errors and assess limits of agreement (mean diff ± 1.96 SD) between the two methods of GA calculation (IVF and CRL reference chart). The narrower these limits are, the better the agreement between the methods. As the data were not Gaussian, the median difference was evaluated by the nonparametric Wilcoxon matched-pairs signed-rank test.

All statistical analyses were performed using the Statistical Package for the Social Sciences Version 22 (SPSS Inc., Chicago, IL) and StatsDirect version 3 (Altrincham UK).

Results

Totally 217 IVF FTSC examinations were carried out in the 10-year period. Of 217, 178 (82%) fulfilled the inclusion criteria for analysis. These women had: mean (SD) age = 35 (5) years, median (IQR) booking BMI = 25 (23–28) kg/m². Of 178, 177 (99.4%) were non-smokers, and 173/177 (97.7%) were Caucasian.

GA comparisons between IVF date and CRL length reference ranges.

- (1) Systematic differences between IVF GA and USS/CRL reference range GA: When comparing the USS GA (days) calculated from the Robinson & Fleming CRL reference chart, and the 'true' GA (days) calculated from the known fertilization date, our results demonstrate a systematic overestimation with most data points falling below the 0 difference mark (i.e. the point where both GA time periods in days were the same) (Figure 1).

Systematic overestimations were identified from Bland–Altman plots of IVF GA days and the four other reference charts were available.

- (2) Mean differences between IVF GA and USS/CRL reference range GA.

When comparing the USS/Robinson & Fleming CRL reference chart GA, and the 'true' GA calculated from the known fertilization date, the mean difference identified was 3.0 days (95% CI: 2.70 to 3.36 days) ($p < 0.001$).

Systematic overestimations (average range from 1.6 to 2.9 days) were obtained from the four other reference charts available (Table 1).

Discussion

Our study has demonstrated a systematic and significant difference between USS GA calculated from the currently recommended UK CRL reference chart⁸ and a 'true' GA based on known implantation dates in IVF pregnancies.

We acknowledge several limitations in our study: IVF pregnancies may not be biologically equivalent to spontaneous conceptions; however our stringent inclusion/exclusion criteria aimed to provide a homogenous sample population with accurate and known gestation age to enable meaningful comparison between true GA and dates by CRL.

USS scanning was performed by a large number of individuals and we were unable to assess inter and intra rater reliability, but all scan operators are required to be trained to FMF and DQASS standard.⁶

Our cohort was predominately Caucasian which while limiting the generalizability of our findings reduces the impact of potential confounding of our results from ethnic diversity.

CRL is known to increase with increasing maternal age.¹⁶ The currently recommended Robinson and Fleming CRL reference chart was developed in 1975 when the average maternal age was 26 years while in our cohort the average maternal age was 35 years. This may be a confounding factor which we were unable to

Figure 1. Bland–Altman plot comparing IVF GA and Robinson & Fleming CRL reference chart GA.

Table 1. Comparison of CRL reference range GA with known IVF date.^a

CRL reference chart	Mean [95% CI] difference from IVF date (days)	Limits of agreement	<i>p</i> value
Robinson and Fleming ⁸	−3.0 [−3.34 to −2.69]	−7.34 to 1.31	<0.0001
McLennan and Shulter ⁹	−1.65 [−1.96 to −1.35]	−5.68 to 2.38	<0.0001
Sahota et al. ¹⁰	−2.90 [−3.22 to −2.58]	−7.13 to 1.33	<0.0001
Verburg et al. ¹¹	−2.31 [−2.62 to −2.0]	−6.48 to 1.86	<0.0001
Papageorgiou et al. ¹²	−2.74 [−3.09 to −2.39]	−7.35 to 1.87	<0.0001

CRL: crown rump length; GA: gestational age.

^aGA = IVF days − CRL days.

correct for in the present study, however, in our dataset, there was no significant correlation between maternal age and CRL, $r = 0.128$, $p = 0.09$.

Despite these limitations, our study has several strengths: were able to utilize complete data sets on a relatively large, well defined population to allow meaningful comparison of the two methods of GA estimation. Routine ultrasound examinations are now carried out by well-trained specialist sonographers with up-to-date high resolution ultrasound equipment which were not available when the Robinson and Fleming CRL reference range was developed.

We are unable to account for the systematic overestimation seen with the reference ranges developed following these improvements in scanning technology and sonographer training: the ‘best’ performing alternative reference range (i.e. having the lowest average overestimation) still systematically overestimates by 1.65 days.

The increasing importance of identifying the at-risk foetus as early as possible has directly impacted on service provision, with increasing numbers of women referred for extra scans due to perceived reduced foetal growth based on deviations from the individualized foetal growth trajectories derived from the FTSC USS. While an overestimation of three days may appear to be of limited clinical impact, this could have important implications for maternity care, particularly if we wrongly identify a foetus as being small for gestational age (SGA) following routine assessment at 20 weeks gestation: an assessment based on alternative foetal biometry reference charts (abdominal circumference (AC), femur length (FL) and head circumference (HC)) utilizing a potentially inaccurate first trimester CRL. This could subsequently result in potentially unnecessary interventions and pre-term delivery as per Royal College of Obstetricians and Gynaecologists (RCOG) guidelines.¹⁷

Conclusion

The current CRL reference chart used locally in routine clinical practice would appear to systematically overestimate GA by an average of three days when assessed in IVF pregnancies. Systematic overestimations were also identified in alternative available reference charts. While these differences may appear slight, the systematic error we identified may have potential implications on the accuracy of GA estimation in routine antenatal care, particularly when related to those pregnancies at risk of pre-term delivery or growth restriction. Further research is warranted to confirm our findings in larger, non IVF cohorts which could lead to the need for a more accurate CRL reference range.

Acknowledgments

Obstetric Ultrasound Team RD&E NHS FT, Exeter Fertility, RD&E NHS FT and NIHR Exeter Clinical Research Facility.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) declared receipt of the following financial support for the research, authorship, and/or publication of this article: This article presents independent research funded by RD&E NHS FT Research and Development Department Small Grant Scheme supported by the National Institute for Health Research (NIHR) Exeter Clinical Research Facility. The views expressed are those of the author(s) and not necessarily those of the RD&E, NHS, the NIHR or the Department of Health.

Ethical approval

Ethical approval was given by the Health Research Authority (HRA): IRAS Project ID 206176, REC reference 16/HRA/1424.

Guarantor

CI Dr Helen Liversedge

Contributors

All authors contributed to the development of the project. BK was responsible for obtaining governance approvals. AB was responsible for data collection. BK, RP and HL were responsible for data analysis. BK wrote the original draft of the paper. AB, RP, HL reviewed and revised the original draft. All authors agreed the final version to be published and are accountable for this work.

References

1. Tunon K, Eik-Nes SH, Grottnum P, et al. Gestational age in pregnancies conceived after *in vitro* fertilization: a comparison between age assessed from oocyte retrieval, crown-rump length and biparietal diameter. *Ultrasound Obstet Gynecol* 2000; 15: 41–46.
2. Wald NJ, Cuckle HS, Densm JW, et al. Maternal serum screening for Down's syndrome: the effect of routine ultrasound scan determination of gestational age and adjustment for maternal weight. *Br Journal Obstet Gynaecol* 1992; 99: 144–149.
3. Blondel B, Morin I, Platt RW, et al. Algorithms for combining menstrual and ultrasound estimates of gestational age: consequences for rates of preterm and postterm birth. *BJOG* 2002; 109: 718–720.
4. Reuss ML, Hatch MC and Susser M. Early ultrasound dating of pregnancy: selection and measurement biases. *J Clin Epidemiol* 1995; 48: 667–674.
5. Taipale P and Hiilesmaa V. Predicting delivery date by ultrasound and last menstrual period in early gestation. *Obstet Gynecol* 2001; 97: 189–194.
6. Fetal Anomaly Screening Programme. *Handbook for ultrasound practitioners*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/443865/FASP_ultrasound_handbook_July_2015_090715.pdf (2015 accessed 13 March 2017).
7. Napolitano R, Dhama J, Ohuma EO, et al. Pregnancy dating by fetal crown rump length: a systematic review of charts. *Br J Obstet Gynaecol* 2014; 21: 556–565.
8. Robinson HP and Fleming JEE. A critical evaluation of sonar 'crown rump length' measurements. *Br J Obstet Gynaecol* 1975; 82: 702–710.
9. McLennan AC and Schluter PJ. Construction of modern Australian first trimester ultrasound dating and growth charts. *J Med Imaging Radiat Oncol* 2008; 52: 471–479.
10. Sahota DS, Leung TY, Leung TN, et al. Fetal crown-rump length and estimation of gestational age in an ethnic Chinese population. *Ultrasound Obstet Gynecol* 2009; 33: 157–160.
11. Verburg BO, Steegers EAP, Ridder MDE, et al. New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from population-based cohort study. *Ultrasound Obstet Gynecol* 2008; 31: 388–396.
12. Papageorghiou AT, Kennedy SH, Salomon LJ, et al. International standards for early size and pregnancy dating based on ultrasound measurement of crown-rump length in the first trimester pregnancy. *Ultrasound Obstet Gynecol* 2014; 44: 641–648.
13. Loughna P, Chitty L, Evans T, et al. Fetal size and dating: charts recommended for clinical obstetric practice. *Ultrasound* 2009; 17: 161–167.
14. GE Healthcare. ViewPoint 6 for OB/GYN, Available at: http://www3.gehealthcare.com/en/products/categories/ultrasound/ultrasound_it/viewpoint_for_ob-gyn (accessed 13 March 2017).
15. Bland JM and Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307–310.
16. Bottomley C, Daeman A, Mukri F, et al. Assessing first trimester growth: the influence of ethnic background and maternal age. *Hum Reprod* 2009; 24: 284–290.
17. Royal College of Obstetricians and Gynecologists. The investigation and management of the small-for-gestational-age fetus. Available at: https://www.rcog.org.uk/globalassets/documents/guidelines/gtg_31.pdf (2013, accessed 30 August 2017).