Reproducibility of first trimester placental volume assessment with fully automated placental segmentation using the OxNNet toolkit

Sam Mathewlynn*†, Yi Yin*, Theophilus Adu-Bredu*, Mohammadreza Soltaninejad*, Katja Shipp*, Jess Farmer*, Debbie Hedgecott*, Sally Collins*‡

* Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK
† Oxford University Hospitals NHS Foundation Trust, Oxford, UK
‡ Birmingham Women's Hospital, Birmingham, UK

BACKGROUND

- First trimester placental volume (PV) shows potential as a predictor of fetal growth restriction (FGR) and pre-eclampsia (PE), but large-scale research and clinical utility have been hindered by the time-consuming process of 3D placental segmentation. [1–5]
- OxNNet is a novel deep learning tool for rapid and fully automated placental segmentation and volume calculation, built using fully convolutional artificial neural network architecture.
- The ongoing *Oxford Placental Ultrasound Study* (OxPLUS) will recruit a cohort of 3,500 pregnancies in which 3D placental ultrasound is performed in the first trimester, with images analysed using OxNNet. Novel predictive models (which will include OxNNet derived PV) for FGR and PE will be developed and evaluated.
- For these models to be useful, it is critical to understand the extent to which PV measures are reproducible.

OxPLUS

Scan the QR code for more information about the Oxford Placental Ultrasound Study

METHODS

- Participants opportunistically selected from the OxPLUS cohort underwent 3D placental ultrasound between 11^{+2}–14^{+1} weeks' gestation. The placenta was segmented, and the PV calculated, using OxNNet.
- Assessments were either performed twice by Operator 1, or once each by Operator 1 and Operator 2. The ultrasound machine was reset between assessments and assessors were blinded to each other's performance.
- All analyses were performed using R (version 4.3.1).
- Sample size was calculated according to Bland's (2009) recommendation (Rho 0.8, confidence width 0.2, confidence level 0.95). [6]
- Bland-Altman plots were constructed with accompanying descriptive statistics using mean difference ± 1.96 standard deviations (SD) as the limit of agreement.
- The data were not normally distributed; therefore, transformation was performed according to the *bestNormalize* package (square root and inverse hyperbolic sine for intra- and inter-operator datasets, respectively)
- Intraclass correlation coefficients (ICC) with 95% confidence intervals.

AIM

To evaluate the intra-operator and inter-operator reproducibility of placental volume measurements derived using the OxNNet toolkit.

FIGURE 1

Bland-Altman plot for placental volume intra-operator testing

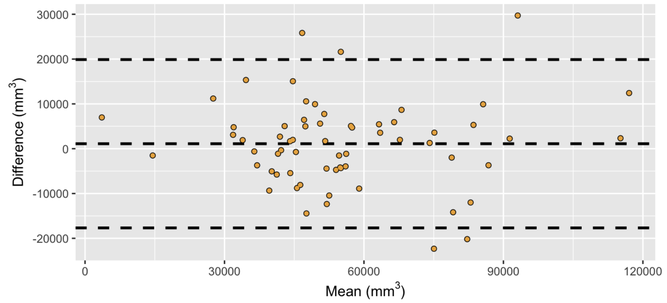
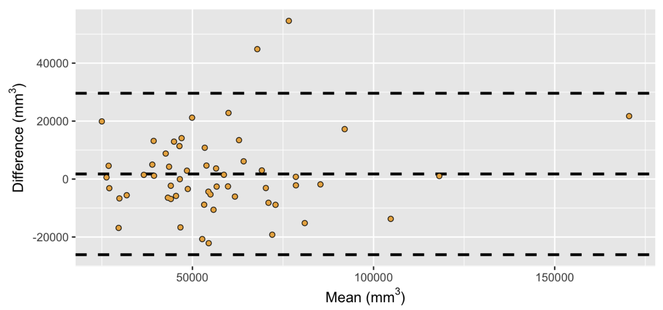


FIGURE 2

Bland-Altman plot for placental volume inter-operator testing



INTERPRETATION OF ICC

There are no agreed standards for what constitutes an acceptable ICC. The threshold of acceptability will vary according to the context in which a measure is to be used, and the width of the 95% confidence interval should be considered. However, ICC values greater than 0.5, 0.75, and 0.9 generally suggest moderate, good, and excellent reliability, respectively. [7]

RESULTS

- The cohorts included 64 and 56 cases for intra-operator and inter-operator assessment, respectively.
- The intra-operator ICC was 0.90 (95% CI 0.84–0.94) and the mean difference was 1,110 mm^3. 92.2% (n=59) of cases were within agreement limits (Table 1).
- The inter-operator ICC was 0.79 (95% CI 0.66–0.87) and the mean difference was 1,750 mm^3. 96.4% (n=54) of cases were within agreement limits (Table 1).
- Bland-Altman plots are shown in Figures 1-2.

TABLE 1

Results of intra-operator and inter-operator reliability testing

	Intra-operator N=64	**Inter-operator** N=56
ICC with 95% CI	0.90 (0.84–0.94)	0.79 (0.66–0.87)
Mean difference (mm^3)	1,110	1,750
Agreement limits (mm^3)	-17,670–19,890	-26,100–29,559
Within agreement limit (%)	92.2	96.4
Above agreement limit (%)	4.7	3.6
Below agreement limit (%)	3.1	0.0

CONCLUSION

First trimester placental volume measurements derived from 3D ultrasound images using OxNNet demonstrate excellent intra-operator reproducibility and good inter-operator reproducibility.

REFERENCES

1. Hafner E, Metzenbauer M, Dillinger-Paller B, Hoefinger D, Schuchter K, Sommer-Wagner H, Philipp K. Correlation of first trimester placental volume and second trimester uterine artery Doppler flow. Placenta. 2001 Sep-Oct;22(8-9):729-34
2. Plasencia W, Akolekar R, Dagklis T, Veduta A, Nicolaides KH. Placental volume at 11-13 weeks' gestation in the prediction of birth weight percentile. Fetal Diagn Ther. 2011;30(1):23-8.
3. Law LW, Leung TY, Sahota DS, Chan LW, Fung TY, Lau TK. Which ultrasound or biochemical markers are independent predictors of small-for-gestational age? Ultrasound Obstet Gynecol. 2009 Sep;34(3):283-7. doi: 10.1002/uog.6455. PMID: 19670336.
4. Hafner E, Metzenbauer M, Höfinger D, Stonek F, Schuchter K, Waldhör T, Philipp K. Comparison between three-dimensional placental volume at 12 weeks and uterine artery impedance/notching at 22 weeks in screening for pregnancy-induced hypertension, pre-eclampsia and fetal growth restriction in a low-risk population. Ultrasound Obstet Gynecol. 2006 Jun;27(6):652-7.
5. Hashish N, Hassan A, El-Semary A, Gohar R, Youssef MA. Could 3D placental volume and perfusion indices measured at 11-14 weeks predict occurrence of preeclampsia in high-risk pregnant women? J Matern Fetal Neonatal Med. 2015 Jun;28(9):1094-8.
6. Bland JM. The tyranny of power: is there a better way to calculate sample size? BMJ. 2009 Oct 6;339:b3985.
7. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016 Jun;15(2):155-63

Corresponding author: sam.mathewlynn@wrh.ox.ac.uk