7-1-2017

Comparison of modified two-point dixon and chemical shift encoded MRI water-fat separation methods for fetal fat quantification: Compare Water-Fat Methods for Fetal Fat

Stephanie A. Giza

Michael R. Miller

Prasiddha Parthasarathy

Barbra de Vrijer

Charles A. McKenzie

Title: Comparison of modified two-point Dixon and IDEAL water-fat separation methods for fetal fat quantification

Abstract

Background: Fetal fat is indicative of the energy balance within the fetus, which may be disrupted in pregnancy complications such as fetal growth restriction, macrosomia, and gestational diabetes. Water-fat separated MRI is a technique sensitive to tissue lipid content, measured as fat fraction (FF), and can be used to accurately measure fat volumes. Modified two-point Dixon and IDEAL are water-fat separated MRI techniques that could be applied to imaging of fetal fat.

Purpose/Hypothesis: Here we compare the methods, as modified two-point Dixon has biases present that are corrected in IDEAL which may contribute to differences in the measurement of fetal fat volume and FF.

Study Type: Cross-sectional study for comparison of two MRI pulse sequences.

Population/Subjects/Phantom/Specimen/Animal Model: Twenty-one pregnant women with singleton pregnancies.

Field Strength/Sequence: 1.5T, modified two-point Dixon and IDEAL.

Assessment: Manual segmentation (A1) of total fetal fat volume and mean FF from modified 2-point Dixon and IDEAL FF images.

Statistical Tests: Reliability was assessed by calculating the intraclass correlation coefficient (ICC). Agreement was assessed using a one-sample t-test on the fat measurements difference values (modified two-point Dixon - IDEAL) for both fat volume and mean FF. The difference

scores were tested against a value of 0, which would indicate that the measurements were identical.

Results: The fat volume and FF measured by modified two-point Dixon and IDEAL had excellent reliability, demonstrated by ICCs of 0.93 (p<0.001) and 0.90 (p<0.001) respectively. They were not in agreement, with IDEAL giving mean fat volumes 180 mL greater and mean FF 3.0% smaller than modified two-point Dixon.

Data Conclusion: The reliability between modified two-point Dixon and IDEAL indicates that either technique can be used to compare fetal fat measurements in different participants, but they are not in agreement due to uncorrected biases in modified two-point Dixon.

Keywords

Fetal

Water-Fat MRI

Adipose Tissue

IDEAL

Dixon

Fat Quantification

Introduction

The assessment of abnormalities in fetal fat development may provide insight into fetal metabolic health because it is reflective of the energy balance within the fetus (1). This energy balance may be disrupted in the case of placental insufficiency or when there are disruptions in maternal metabolism (obesity, diabetes) (2). Previous imaging studies have found increased amounts of adipose tissue in macrosomic fetuses (3-5) and fetuses of diabetic mothers (6-8), and decreased adipose tissue in fetal growth restriction fetuses (8-10). However, there are limited imaging studies that have examined the lipid content of the fetal fat, which changes through gestation as the adipocytes are maturing.

Water-fat MRI is an ideal technique to assess fetal fat development as it is sensitive to the amount of lipid within a tissue (11,12). Two-point Dixon and IDEAL are water-fat MRI methods that can be used to assess the lipid content of fetal adipose tissue. Each of these techniques have different strengths and weaknesses, which should be compared for the application of assessing fetal fat development.

The modified two-point Dixon technique uses an opposed-phase gradient recalled echo combined with two-point Dixon water-fat separation (13). Using two echoes, two-point Dixon produces in-phase and opposed-phase images, which are added or subtracted to give fat-only and water-only images. These can then be used to produce a FSF (fat/(water + fat)) image. Like two-point Dixon, IDEAL is a gradient recalled echo sequence, but instead of two echoes it acquires six echoes that are used for water-fat separation (14-16). The addition of multiple echoes allows for modelling of more variables than that of modified two-point Dixon. In IDEAL, the fat

fraction values calculated are equal to proton-density fat fraction (PDFF) because of the additional variables (14-16).

Both techniques correct for B_0 field inhomogeneities; in modified two-point Dixon this is done through post-processing using a region-growing phase-correction algorithm developed by Ma (13). In IDEAL, the B_0 field inhomogeneities are accounted for in the model used to fit the data (16). Both modified two-point Dixon and IDEAL can use small flip angles to reduce the effects of T1 relaxation and therefore minimise T1 bias (17).

IDEAL differs from modified two-point Dixon in the modelling of both R2* relaxation and a six-peak fat spectrum to the acquired data (13-16). Meisamy et al. showed through comparison to magnetic resonance spectroscopy that failing to correct for R2* relaxation and a multi-peak fat spectrum introduces a bias in the fat fraction measured in adult liver (15). Since these biases are not corrected in modified two-point Dixon, the FSF measured may not be as accurate as the PDFF obtained through IDEAL.

The strength of modified two-point Dixon lies in its availability and speed. Modified two-point Dixon is a more widely available sequence than IDEAL, which increases the use of this technique in multiple institutions. IDEAL requires the acquisition of images at more echo times (TEs) than modified two-point Dixon (at least 6 echoes, vs. 2), resulting in inherently longer acquisitions for the same resolution and anatomic coverage. Additionally, IDEAL has a lower signal to noise ratio (SNR) per unit time as compared to modified two-point Dixon (13,18). For

this reason, modified two-point Dixon would allow the acquisition of images with better SNR in a shorter amount of time, which is key in fetal imaging where motion is a major concern.

The purpose of this study was to compare modified two-point Dixon and IDEAL techniques for the quantification of fetal adipose tissue volume and PDFF/FSF.

Materials and Methods

This study was approved by our institution's Human Studies Research Ethics Board (REB# xxxxxx). Informed consent was obtained when women were recruited from low risk and specialized high body mass index (BMI) obstetric clinics at our institution. Inclusion criteria consisted of the following: pregnant women over the age of 18 between 29 and 38 weeks gestational age. Patients with any medical contraindication to safely undergoing a non-contrast MRI, weight/body habitus that would prevent a successful MRI study, or multiple pregnancy were excluded. One participant was excluded due to medical contraindication to safely undergoing a non-contrast MRI. Clinical data were collected from the participant's charts including gestational age at time of MRI, pre-pregnancy BMI, and diagnoses of fetal growth restriction or maternal diabetes (Table 1).

Consenting participants underwent a fetal MRI in a wide-bore (70 cm diameter) 1.5T MRI (General Electric Optima 450w, Milwaukee, WI, USA) with a 32-coil abdominal phased array. Women were positioned left decubitus or rolled towards a left lateral decubitus position with a cushion under either their back or their right side for comfort. Scout images (T2 weighted Single Shot Fast Spin Echo (SSFSE)) were acquired to locate the fetus and determine its orientation.

Both 3D IDEAL (specific implementation IDEAL IQ) and modified 3D two-point Dixon (specific implementation LAVA-Flex) volumes were acquired in a plane axial to the fetal abdomen during a maternal breath hold (imaging parameters in Table 2). The two-point Dixon and IDEAL volumes were prescribed to match the anatomic coverage as closely as possible. Phase FOV and resolution, and slice thickness were altered as necessary to allow acquisition in the maternal breath hold. A second 3D IDEAL volume with anatomic coverage matched to the first acquisition was acquired at the end of each MRI exam.

Water-only, fat-only, PDFF and R2* maps were reconstructed from the 3D IDEAL data using the method of Yu et al. (16,18,19). Water-only and fat-only images were reconstructed from the 3D Dixon data using the method of Ma et al. (13). FSF image volumes were calculated voxel-by-voxel from the two-point Dixon fat and water signal intensities (FSF = Fat/(Water+Fat)).

Total fetal fat from the entire fetal volume was manually segmented (A1, 2 years experience) from all the PDFF/FSF images using 3D Slicer (4.7.0 nightly build 2016-12-06)(20-22) by tracing the along the border of high signal intensity corresponding to fetal fat. This included subcutaneous fat, perirenal fat, orbital fat, paravertebral fat and bone marrow when visible. Measured variables included the total segmented volume and the mean PDFF/FSF within the segmented volume.

To investigate the failure to correct for R2* decay resulting in the appearance of artefactual "pseudo-fat" in Dixon acquisitions, all modified two-point Dixon volumes were evaluated for the appearance of fetal liver fat through observation by two trained readers (A1 2 years experience, A5 25 years experience). The image volumes were sorted into either a pseudo-fat group when fat appeared in the fetal liver, or a pseudo-fat-free group, when no fat appeared in the fetal liver. The occurrence of elevated R2* in the fetal liver was evaluated by placing a 15mm spherical ROI

within the fetal liver. R2* and PDFF values were measured from these liver ROIs on IDEAL images.

The reliability of fat volume and PDFF/FSF measurements made with modified two-point Dixon and IDEAL were tested to assess the consistency of ordering between participants between the modified two-point Dixon and IDEAL. The inter-rater reliability of fat volume and PDFF were tested to assess the consistency of ordering between participants between two trained independent readers (A1 2 years experience, A3 no previous experience), who each manually segmented 10 IDEAL PDFF sets. The inter-rater reliability of modified two-point Dixon segmentation was previously found to be very high (23). Test-retest reliability of IDEAL derived fat volume and PDFF measurements were assessed by comparing measurements from the two 3D IDEAL acquisitions to assess the consistency of ordering between participants for each test. Intra-class correlation was used to test all three reliability assessments.

To assess the level of agreement (i.e., identical results) between modified two-point Dixon and IDEAL measurement of fat volume and PDFF/FSF, a one-sample t-test was performed on the fat measurements difference values (modified two-point Dixon minus IDEAL). These difference scores were tested against a value of 0, which would indicate that the measurements were identical. The average modified two-point Dixon/IDEAL measurement was plotted against the difference between the methods in Bland-Altman plots. The level of agreement for inter-rater and test-retest was examined in the same manner, with Bland-Altman plots using log values when appropriate. Mann-Whitney U tests were applied to detect differences between the pseudo-fat and pseudo-fat free groups' R2*, PDFF, and gestational ages. All statistical analyses were conducted in SPSS v.24 (IBM Corporation, Armonk, NY, USA), and p-values < 0.05 were considered statistically significant.

Results

Twenty-nine women participated in the study. One participant was unable to complete the MRI, 3 were excluded from analysis for fetal motion, and 2 were excluded from analysis for severe artefact. Two fetuses had no segmentable fat. Of the remaining 21 fetuses, one had severe motion artefacts in modified two-point Dixon, leaving 20 fetuses for the comparison of modified two-point Dixon and IDEAL. 3D renderings and segmentations of total fetal fat from modified two-point Dixon and IDEAL are shown in Figure 1. The ICC between modified two-point Dixon and IDEAL for fetal fat volume was 0.928 (p<0.001) and for fetal PDFF/FSF was 0.898 (p<0.001), indicating strong reliability for both measurements. The t-test for fetal fat volume difference (modified two-point Dixon – IDEAL) against 0 was significant (p<0.001), indicating that the two methods were not statistically identical. The mean and standard deviation for fetal fat volume measured by modified two-point Dixon was 980 ± 360 mL, and by IDEAL was 1180 ± 390 mL. The mean difference (modified two-point Dixon – IDEAL) for fetal fat volume was -180 mL. The Bland-Altman plot for fetal fat volume is shown in Figure 2A. The t-test for fetal PDFF/FSF difference (modified two-point Dixon – IDEAL) against 0 was also significant (p=0.001), again indicating that the two methods were not statistically identical. The mean difference (modified two-point Dixon – IDEAL) for PDFF/FSF was 3.0%. The Bland-Altman plot for fetal PDFF/FSF is shown in Figure 2B.

7/21 fetuses had pseudo-fat appear in their livers with modified two-point Dixon, while no fetuses had detectable liver fat with IDEAL. Sample images including modified two-point Dixon fat images, IDEAL PDFF images, and IDEAL R2* maps for pseudo-fat and pseudo-fat free groups are shown in Figure 3. Mann-Whitney U tests indicated that there was a significant

difference in the fetal liver R2* between the pseudo-fat and pseudo-fat free groups (p<0.001) (Figure 4A), but not in their liver PDFF measured by IDEAL (p=0.332) (Figure 4B) or their gestational ages at MRI (p=0.654) (Figure 4C). All the fetuses with pseudo-fat had liver R2* values ≥ 45 s^{-1}, whereas all the fetuses without pseudo-fat had liver R2* values ≤ 35 s^{-1}.

17 fetuses had 2 sets of motion free IDEAL images and were used for test-retest reliability. The ICC for fetal fat volume was 0.971 (p<0.001) and for fetal PDFF was 0.980 (p<0.001), indicating strong test-retest reliability. The t-test for difference (test-retest) against 0 for fetal fat volume was not significant (p=0.152), indicating the results were statistically similar. No proportional bias was detected as shown in the Bland Altman plot (Figure 5A). The t-test for difference (test – retest) against 0 for fetal PDFF was significant (p=0.045), indicating that the test-retest were not statistically identical. The magnitude of the difference gives an estimate of the bias detected, which was 0.9%. The Bland-Altman plot is shown in Figure 5B.

The ICC between the two readers for fetal fat volume was 0.897 (p=0.002), and for fetal PDFF was 0.946 (p<0.001). This shows a strong inter-rater reliability. The t-tests for difference (reader 1 – reader 2) against 0 for fetal fat volume and PDFF were not significant (p=0.847 and p=0.706), indicating the results were statistically similar. No proportional bias was detected as shown in Bland-Altman plots (Figure 6).

Discussion

Our results indicate that modified two-point Dixon and IDEAL measurements of fat fraction and fat volume are reliable, but not in agreement. Although both methodologies have good ICCs and

can be used for comparison between participants, differences between measured total fetal fraction warrants questions regarding accuracy of one or both methods.

Modified two-point Dixon has biases present in the FSF because it does not correct for R2* and does not use a multipeak fat spectrum (13). Meisamy et al. have shown that failing to incorporate these two factors into the modeling of signal from Dixon methods introduces a bias compared to spectroscopy results (15).

In our imaging of fetal fat, we observed examples of this bias in the fetal liver. When the fetal liver had elevated R2* (short T2*), pseudo-fat appeared in the liver on modified two-point Dixon fat images. This was not seen in the IDEAL images, since R2* is accounted for in the modelling. This was a common occurrence, with one third of the fetuses we imaged having appearance of pseudo-fat in the liver on modified two-point Dixon images. It has previously been shown that the fetal liver in the third trimester has a longer R2* (50 s^{-1} at 1.5T) compared to that in the adult liver (36 s^{-1} at 1.5T) (24). This occurs because the fetal liver is a major hemopoietic site during fetal development and therefore acts as a reservoir for iron in the third trimester (24). R2*-weighted gradient recalled echo sequences have previously been used in the investigation of hemochromatosis (25), and together these results suggest that IDEAL could be used in the future to investigate hematopoiesis through gestation as well as prenatal detection of hemochromatosis.

We believe that IDEAL is the preferred method for measuring fetal fat fraction, as the biases present in modified two-point Dixon have been corrected (16,18,19). This has been shown in the adult liver by comparing to spectroscopy (15), and it is likely that similar factors are at play in the fetus, as demonstrated in the fetal livers from our study.

Previous studies have used PDFF to distinguish brown adipose tissue from white adipose tissue in infants, as brown adipose tissue has a lower PDFF (26-28) Unbiased fat fraction measurements are key to this differentiation, therefore IDEAL would be best suited to an investigation of brown adipose tissue in the fetus. This would be especially important for assessing brown adipose tissue across gestation, as white adipose tissue has a lower lipid content at earlier gestational ages (29), and therefore the difference between the fat fractions of brown and white adipose tissue may be small.

Test-retest showed that IDEAL is a reliable method for measurement of fetal fat volumes and PDFF. The volumes measured with the IDEAL test-retest agreed; however, their PDFF values were not statistically identical. The mean bias between the two methods was less than 1%, which is small enough that there should be no practical difference in the values measured. The inter-rater reliability and agreement of the fetal fat volumes and PDFF was also good, indicating that this is a robust method for assessment of fetal fat.

Strengths of our study include using 3D techniques, the measurement of total fetal fat and inclusion of a heterogeneous group of participants. By using 3D techniques to measure total fetal fat with both sequences, we minimized the effect of fetal position. Since the fetus is liable to move between sequences, it is possible it is in a different position during modified two-point Dixon and IDEAL acquisitions. This positional effect is reduced by the assessment of the entire fetus with a 3D acquisition, resulting in measuring total fetal fat, regardless of fetal position. We recruited a heterogeneous population of participants, including normal and high BMI, growth restricted fetuses, and diabetic mothers and have demonstrated that our results are applicable to most obstetric populations.

Limitations of this study include testing at one field strength only, limited gestational age range, and uncertainty about the fetal fat spectrum. We performed this comparison at 1.5 T only and although it is expected that results will remain similar at 3.0 T, future studies should be conducted for confirmation. Additionally, future studies at an earlier gestational age than 29-38 weeks can elucidate whether the performance of two-point Dixon and IDEAL are affected by the lower FSF expected earlier in gestation. The expectation is that at lower FSF the biases in two-point Dixon are minimized, whereas IDEAL will struggle at low PDFF values because of its lower signal to noise efficiency. Therefore, it is possible that modified two-point Dixon is preferable at lower gestational ages. Finally, we did not examine the fat spectrum of fetal fat, and it is possible that it differs from the adult fat spectrum employed by Quantitative IDEAL, potentially introducing a source of bias in the fetal PDFF measurements. It has been shown that over the biologically possible range of multipeak fat spectral models there is minimal difference (<2%) when compared in the livers of patients with non-alcoholic steatohepatitis (30), so while it is unlikely that it is different in the fetus this still needs to be examined. Magnetic Resonance Spectroscopy should be used to determine the fetal fat spectrum, investigate gestational age changes and differences from the adult spectrum. This knowledge could be used to alter the spectral model of fat used in IDEAL and allow more accurate PDFF estimation for fetal fat imaging.

In conclusion, either modified two-point Dixon or IDEAL can be used to compare fetal fat volumes and PDFF/FSF between participants. Caution should be used when imaging fetal liver with modified two-point Dixon, particularly in the third trimester where elevated R2* effects are common. In terms of potential biases in measuring PDFF and fat volume in the fetus, we feel that IDEAL is a better method of choice than modified two-point Dixon.

References

1. Toro-Ramos T, Paley C, Pi-Sunyer FX, Gallagher D. Body composition during fetal development and infancy through the age of 5 years. Eur J Clin Nutr 2015;69(12):1279-1289.

2. Ornoy A. Prenatal origin of obesity and their complications: Gestational diabetes, maternal overweight and the paradoxical effects of fetal growth restriction and macrosomia. Reprod Toxicol 2011;32(2):205-212.

3. Hill LM, Guzick D, Boyles D, Merolillo C, Ballone A, Gmiter P. Subcutaneous tissue thickness cannot be used to distinguish abnormalities of fetal growth. Obstet Gynaecol 1992;80(2):268-271.

4. Higgins MF, Russell NM, Mulcahy CH, Coffey M, Foley ME, McAuliffe FM. Fetal anterior abdominal wall thickness in diabetic pregnancy. Eur J Obstet Gynecol Reprod Biol 2008;140(1):43-47.

5. Chauhan SP, West DJ, Scardo JA, Boyd JM, Joiner J, Hendrix NW. Antepartum detection of macrosomic fetus: Clinical versus sonographic, including soft-tissue measurements. Obstet Gynecol 2000;95(5):639-642.

6. Larciprete G, Di Pierro G, Barbati G, et al. Could birthweight prediction models be improved by adding fetal subcutaneous tissue thickness? J Obstet Gynaecol Res 2008;34(1):18-26.

7. Anblagan D, Deshpande R, Jones NW, et al. Measurement of fetal fat in utero in normal and diabetic pregnancies using magnetic resonance imaging. Ultrasound Obstet Gynecol 2013;42(3):335-340.

8. Deans HE, Smith FW, Lloyd DJ, Law AN, Sutherland HW. Fetal fat measurement by magnetic resonance imaging. Br J Radiol 1989;62(739):603-607.

9. Stark DD, McCarthy SM, Filly RA, Callen PW, Hricak H, Parer JT. Intrauterine growth retardation: Evaluation by magnetic resonance. Radiology 1985;155(2):425-427.

10. Larciprete G, Valensise H, Di Pierro G, et al. Intrauterine growth restriction and fetal body composition. Ultrasound Obstet Gynecol 2005;26(3):258-262.

11. Reeder SB, Hu HH, Sirlin CB. Proton density fat-fraction: a standardized MR-based biomarker of tissue fat concentration. J Magn Reson Imaging 2012;36(5):1011-1014.

12. Reeder SB, McKenzie CA, Pineda AR, et al. Water-fat separation with IDEAL gradient-echo imaging. J Magn Reson Imaging 2007;25(3):644-652.

13. Ma J. Breath-hold water and fat imaging using a dual-echo two-point Dixon technique with an efficient and robust phase-correction algorithm. Magn Reson Med 2004;52(2):415-419.

14. Hines CD, Frydrychowicz A, Hamilton G, et al. T(1) independent, T(2) (*) corrected chemical shift based fat-water separation with multi-peak fat spectral modeling is an accurate and precise measure of hepatic steatosis. J Magn Reson Imaging 2011;33(4):873-881.

15. Meisamy S, Hines CDG, Hamilton G, et al. Quantification of Hepatic Steatosis with T1-independent, T2*-corrected MR Imaging with Spectral Modeling of Fat: Blinded Comparison with MR Spectroscopy. Radiology 2011;258(3):767-775.

16. Yu H, Shimakawa A, Hines CD, et al. Combination of complex-based and magnitude-based multiecho water-fat separation for accurate quantification of fat-fraction. Magn Reson Med 2011;66(1):199-206.

17. Liu CY, McKenzie CA, Yu H, Brittain JH, Reeder SB. Fat quantification with IDEAL gradient echo imaging: correction of bias from T(1) and noise. Magn Reson Med 2007;58(2):354-364.

18. Yu H, McKenzie CA, Shimakawa A, et al. Multiecho reconstruction for simultaneous water-fat decomposition and T2* estimation. J Magn Reson Imaging 2007;26(4):1153-1161.

19. Yu H, Shimakawa A, McKenzie CA, Brodsky E, Brittain JH, Reeder SB. Multiecho water-fat separation and simultaneous R2* estimation with multifrequency fat spectrum modeling. Magn Reson Med 2008;60(5):1122-1134.

20. 3D Slicer. Volume 2016; 2017.

21. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 2012;30(9):1323-1341.

22. Kikinis R, Pieper SD, Vosburgh K. 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support. In: Jolesz FA, editor. Intraoperative Imaging Image-Guided Therapy. New York: Springer; 2014.

23. Giza S, Olmstead C, Sinclair K, McKenzie CA, de Vrijer B. Feasibility of Fetal Fat Volume Assessment using 3D Water-Fat MRI. International Society for Magnetic Resonance in Medicine. Singapore; 2016.

24. Goitein O, Eshet Y, Hoffmann C, et al. Fetal liver T2* values: defining a standardized scale. J Magn Reson Imaging 2013;38(6):1342-1345.

25. Marti-Bonmati L, Baamonde A, Poyatos CR, Monteagudo E. Prenatal diagnosis of idiopathic neonatal hemochromatosis with MRI. Abdom Imaging 1994;19(1):55-56.

26. Hu HH, Yin L, Aggabao PC, Perkins TG, Chia JM, Gilsanz V. Comparison of brown and white adipose tissues in infants and children with chemical-shift-encoded water-fat MRI. J Magn Reson Imaging 2013;38(4):885-896.

27. Hu HH, Tovar JP, Pavlova Z, Smith ML, Gilsanz V. Unequivocal identification of brown adipose tissue in a human infant. J Magn Reson Imaging 2012;35(4):938-942.

28. Hu HH, Wu TW, Yin L, et al. MRI detection of brown adipose tissue with low fat content in newborns with hypothermia. Magn Reson Imaging 2014;32(2):107-117.

29. Poissonnet CM, Burdi AR, Garn SM. The chronology of adipose tissue appearance and distribution in the human fetus. Early Hum Dev 1984;10(1-2):1-11.

30. Hong CW, Mamidipalli A, Hooker JC, et al. MRI proton density fat fraction is robust across the biologically plausible range of triglyceride spectra in adults with nonalcoholic steatohepatitis. J Magn Reson Imaging : JMRI 2017.

Tables

Characteristic	n (%)
Maternal BMI (kg/m²)	
Underweight: < 18.5	2 (9.5%)
Normal: 18.5 – 24.9	10 (47.6%)
Overweight: 25 – 29.9	1 (4.8%)
Class I Obesity: 30 – 34.9	1 (4.8%)
Class II Obesity: 35 – 39.9	3 (14.3%)
Class III Obesity: ≥ 40	3 (14.3%)
Diabetic Status	
Pre-existing Type-1 Diabetes	1 (4.8%)
Pre-existing Type-2 Diabetes	1 (4.8%)
Gestational Diabetes	2 (9.5%)
Non-diabetic	17 (81.0%)
Growth Restriction	58.1 (0.5 – 96.6)
Growth Restricted	3 (14.3%)
Appropriate Growth	18 (85.7%)

Table 1. **Participant demographics.** Data is listed as n (%). Total N = 21. BMI = body mass index.

Parameter	Modified two-point Dixon	IDEAL
Repetition Time	6.0 – 6.6 ms	9.7 – 12.7 ms
Flip Angle	5°	6 – 7°
Field of View	50 cm	50 cm
Frequency Encodes	160	128 – 160
Phase Encodes	160	128 – 160
Slice Thickness	4 – 6.5 mm	4 – 6.5 mm
Number of Slices	42 – 64	42 – 78
ARC Acceleration Phase	2x	2x
ARC Acceleration Slice	2x	2.5x
ARC Acceleration Calibration Lines	32x32	32x32
Acquisition Time	10-17 s	12 – 24 s

Table 2. **Imaging Parameters for modified two-point Dixon and IDEAL acquisitions**

Figure Legends

Figure 1. Total fetal fat A) segmentation on modified two-point Dixon, B) segmentation on IDEAL, C) 3D render from modified two-point Dixon, D) 3D render from IDEAL. Images A and B are displayed axial to fetal abdomen through the fetal umbilicus. 3D renders C and D are created from the segmentations in A and B. The hands and feet have limited fat and therefore appear incomplete and patchy in the 3D renders.

Figure 2. Bland-Altman plots of fetal A) fat volume and B) PDFF/FSF from modified two-point Dixon and IDEAL. The solid black line indicates the mean difference between the techniques (A: -180 mL, B: 3.0%), while the two dashed lines indicate the 95% confidence intervals (A: -380, 20 mL, B: -3.6, 9.5%). This demonstrates that modified two-point Dixon underestimates fetal fat volume while overestimating fetal PDFF/FSF compared to IDEAL.

Figure 3. Comparison of pseudo-fat free and pseudo-fat fetuses. A) Modified two-point Dixon fat image without Pseudo-fat in the liver, B) Modified two-point Dixon fat image showing pseudo-fat in the liver, C) IDEAL PDFF image from the same patient and slice as A, D IDEAL PDFF image from the same patient and slice as B, E) IDEAL R2* map from the same patient and slice as A, and F) IDEAL R2* map from the same patient and slice as B. Images are displayed axial to fetal abdomen, with the spherical liver ROI outlined in red. Fat images have been windowed and levelled to display the signal in the fetal livers. The fat fraction measured in image C (pseudo-fat free) was 6.9%, and in image D (pseudo-fat) was 3.7%. The R2* measured in image E (pseudo-fat free) was 31 s^{-1}, and in image F (pseudo-fat) was 45 s^{-1}.

Figure 4. Box and whisker plots of fetal liver A) R2* values and B) PDFF measured from IDEAL, and C) gestational age at MRI for the pseudo-fat and pseudo-fat free groups.

Mann-Whitney U test indicates a significant difference between the groups R2* values (p<0.001), where the pseudo-fat group has a higher fetal liver R2* than the pseudo-fat free group. No significant differences were found in the PDFF or gestational age between the groups (p=0.881, p=0.654). Outliers are shown as circles outside of the box and whisker plots.

Figure 5. Bland-Altman plot of fetal A) fat volume and B) PDFF from IDEAL test-retest. The solid black line indicates the mean difference between the acquisitions (A: -50 mL, B: 0.9%), while the two dashed lines indicate the 95% confidence intervals (A: -340, 230 mL, B: -2.4, 4.2%). This demonstrates that there is no proportional bias in the fetal fat volume between the acquisitions, and a small bias (< 1%) for higher PDFF measurements in first IDEAL acquisition relative to the second acquisition.

Figure 6. Bland-Altman plot of fetal A) fat volume and B) PDFF measured by two readers. The solid black line indicates the mean difference between the two readers' measurements (A: 20 mL, B: -0.4%), while the two dashed lines indicate the 95% confidence intervals (A: -480, 520 mL, B: -7.2, 6.4%). This demonstrates there is no proportional bias between the readers' measurements of fetal fat volume or PDFF.
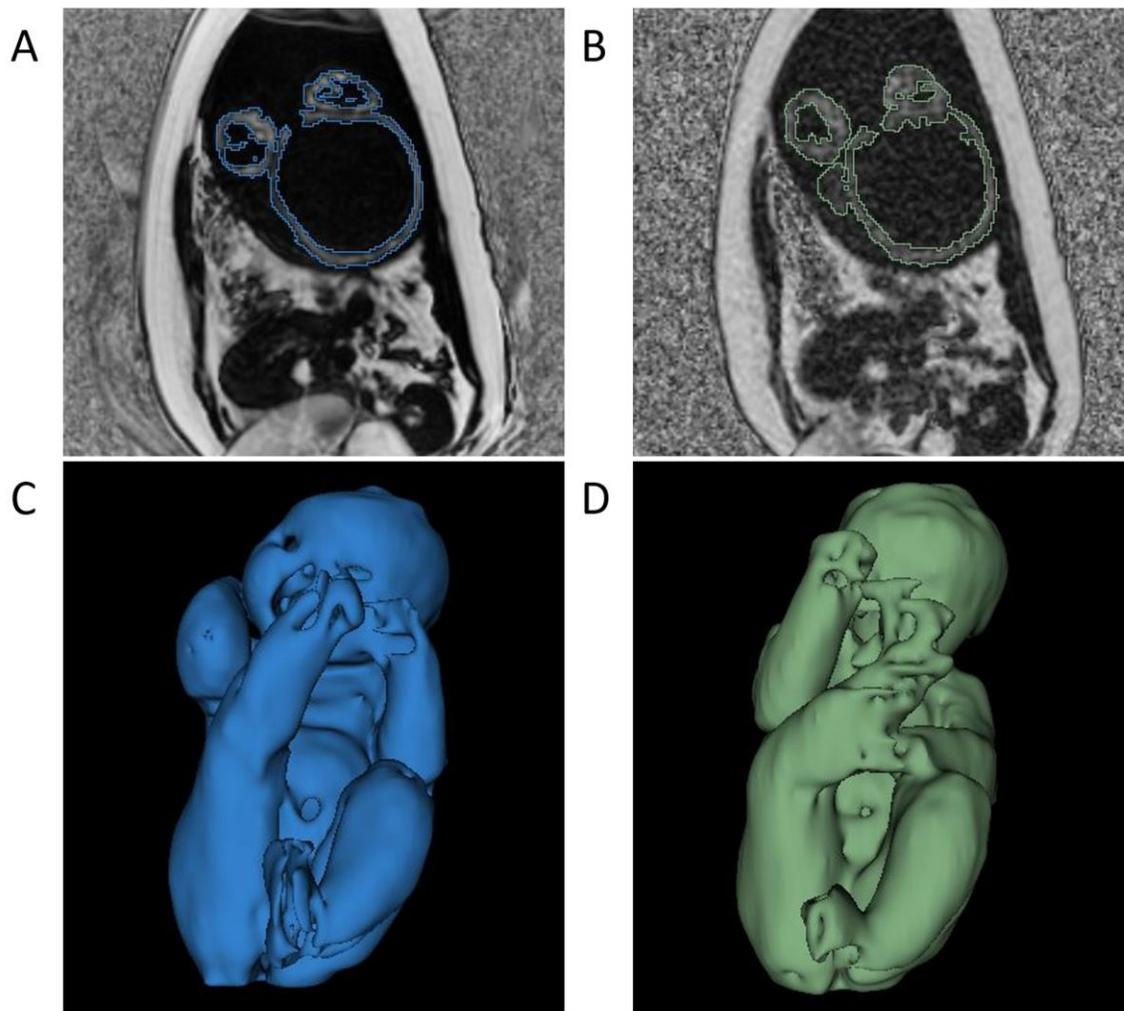
Figures

Figure 1. Total fetal fat A) segmentation on modified two-point Dixon, B) segmentation on IDEAL, C) 3D render from modified two-point Dixon, D) 3D render from IDEAL. Images A and B are displayed axial to fetal abdomen through the fetal umbilicus. 3D renders C and D are created from the segmentations in A and B. The hands and feet have limited fat and therefore appear incomplete and patchy in the 3D renders.
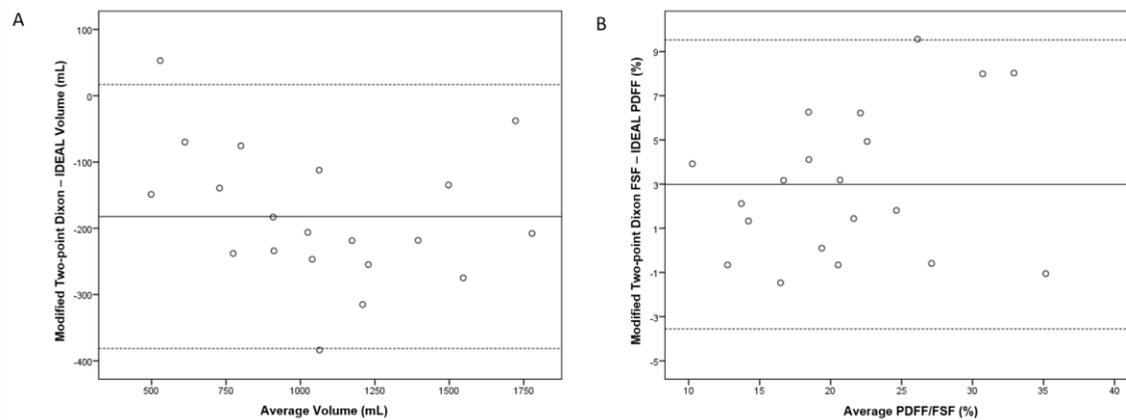
Figure 2. Bland-Altman plots of fetal A) fat volume and B) PDFF/FSF from modified two-point Dixon and IDEAL. The solid black line indicates the mean difference between the techniques (A: -180 mL, B: 3.0%), while the two dashed lines indicate the 95% confidence intervals (A: -380, 20 mL, B: -3.6, 9.5%). This demonstrates that modified two-point Dixon underestimates fetal fat volume while overestimating fetal PDFF/FSF compared to IDEAL.
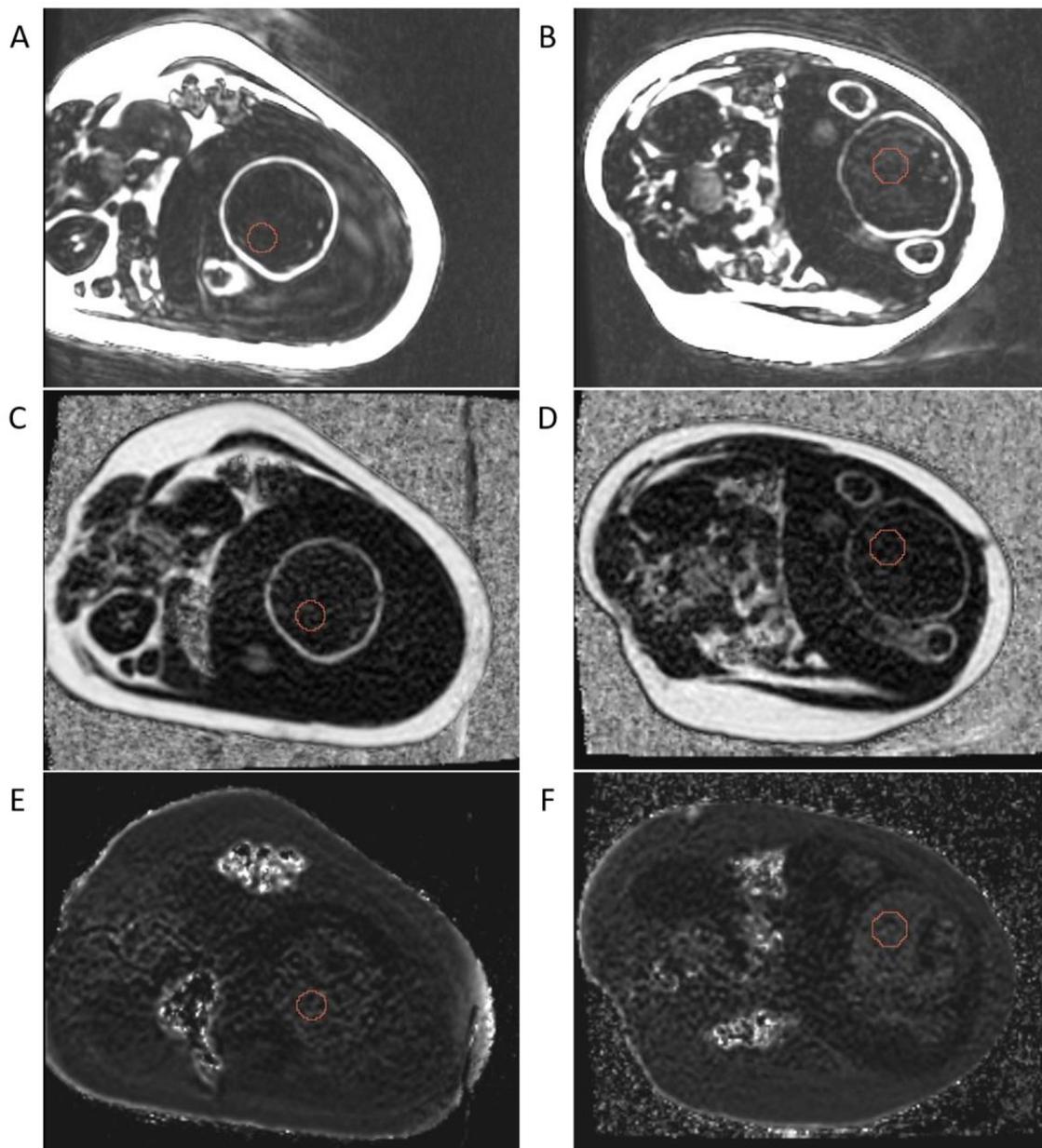
Figure 3. Comparison of pseudo-fat free and pseudo-fat fetuses. A) Modified two-point Dixon fat image without Pseudo-fat in the liver, B) Modified two-point Dixon fat image showing pseudo-fat in the liver, C) IDEAL PDFF image from the same patient and slice as A, D IDEAL PDFF image from the same patient and slice as B, E) IDEAL R2* map from the same patient and slice as A, and F) IDEAL R2* map from the same patient and slice as B. Images are displayed axial to fetal abdomen, with the spherical liver ROI outlined in red. Fat images have

been windowed and levelled to display the signal in the fetal livers. The fat fraction measured in image C (pseudo-fat free) was 6.9%, and in image D (pseudo-fat) was 3.7%. The R2* measured in image E (pseudo-fat free) was 31 s^{-1}, and in image F (pseudo-fat) was 45 s^{-1}.
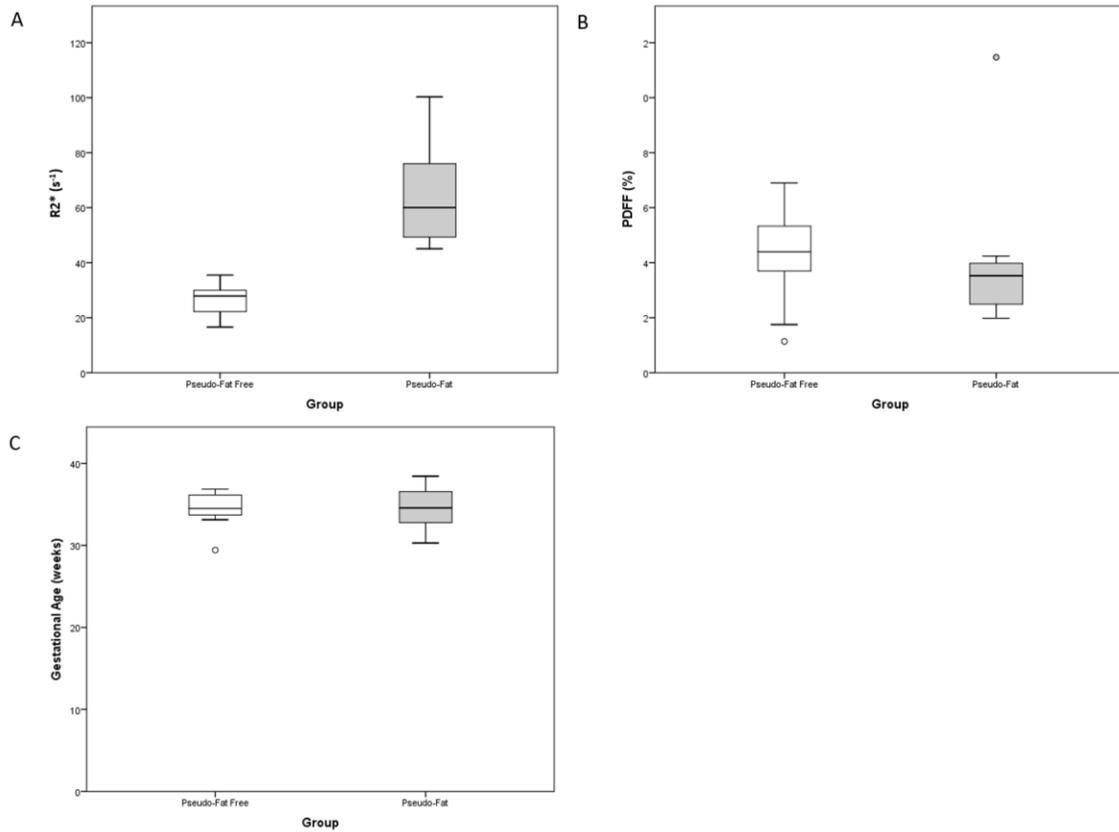
Figure 4. Box and whisker plots of fetal liver A) R2* values and B) PDFF measured from IDEAL, and C) gestational age at MRI for the pseudo-fat and pseudo-fat free groups.
Mann-Whitney U test indicates a significant difference between the groups R2* values (p<0.001), where the pseudo-fat group has a higher fetal liver R2* than the pseudo-fat free group. No significant differences were found in the PDFF or gestational age between the groups (p=0.881, p=0.654). Outliers are shown as circles outside of the box and whisker plots.
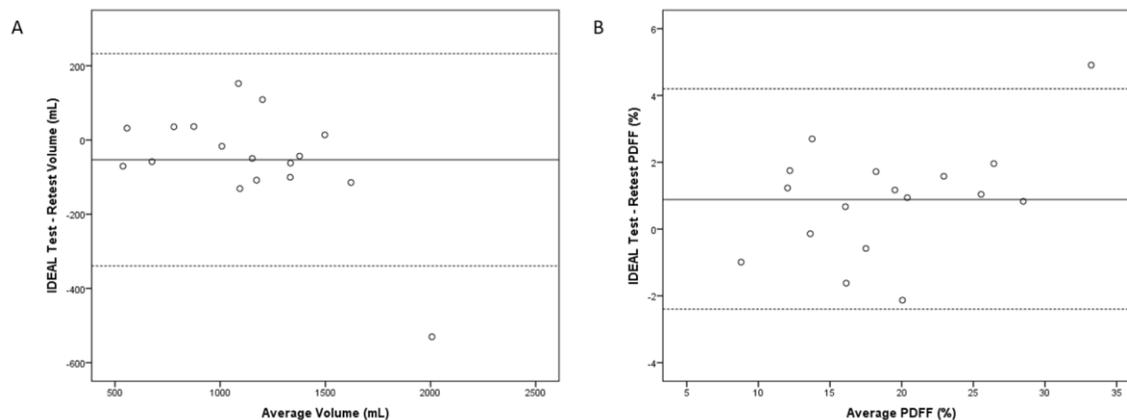
Figure 5. Bland-Altman plot of fetal A) fat volume and B) PDFF from IDEAL test-retest.

The solid black line indicates the mean difference between the acquisitions (A: -50 mL, B: 0.9%), while the two dashed lines indicate the 95% confidence intervals (A: -340, 230 mL, B: -2.4, 4.2%). This demonstrates that there is no proportional bias in the fetal fat volume between the acquisitions, and a small bias ($< 1\%$) for higher PDFF measurements in first IDEAL acquisition relative to the second acquisition.
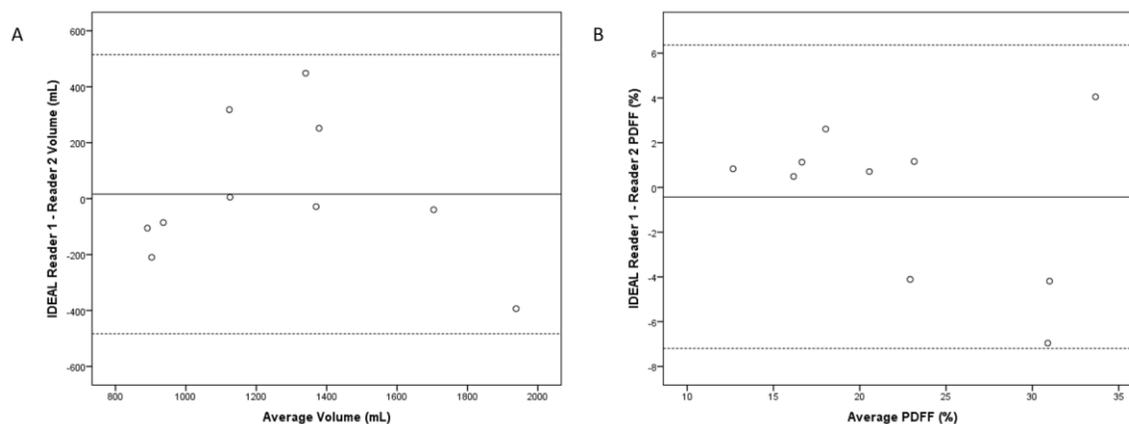
Figure 6. Bland-Altman plot of fetal A) fat volume and B) PDFF measured by two readers.

The solid black line indicates the mean difference between the two readers' measurements (A: 20 mL, B: -0.4%), while the two dashed lines indicate the 95% confidence intervals (A: -480, 520 mL, B: -7.2, 6.4%). This demonstrates there is no proportional bias between the readers' measurements of fetal fat volume or PDFF.