Study	Study groups	Age	Triceps skinfold thickness (mm)				Subscapular skinfold thickness (mm)			
			Controls		IDM		Controls		IDM	
Aman, 2011 [31]	Controls: 28 IDM: 28 (18 T1D, 10 GDM)	<48 hours	5.3 (1.1)		6.7 (1.4)		4.8 (1.1)		6.3 (1.8)	
Brans, 1983 [57]	Controls: 52 IDM: 61 (49 GDM, 12 pre-existing)	<12 hours	Skinfold thickness (sum of triceps and subscapular skinfolds) Control: 7.73 (1.98) IDM 8.87 (1.58)							
Brumbaugh, 2013 [19]	Controls 13 (7 male) IDM 12 (all GDM) (8 male)	16.3±2.3 days [1-3 weeks]	Skinfold thickness (sum of triceps and subscapular skinfolds) Control: 9.88 (1.96) IDM: 11.73 (1.30)							
Brunner, 2013 [28]	Controls: 152 (82 male) IDM: 9 (all GDM) (3 male)	3-5 days	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
			4.6 (0.9)	4.7 (0.8)	5.5 (1.6)	5.0 (0.6)	4.4 (0.9)	4.6 (1.0)	4.9 (1.4)	4.6 (0.3)
			Pooled		**Pooled**		**Pooled**		**Pooled**	
			4.7 (0.9)		5.2 (0.9)		4.5 (1.0)		4.7 (0.8)	
Buhling, 2012 [29]	Controls: 142 IDM: 30 (all GDM)	<72 hours	4.8 (1.5)		4.6 (0.9)		4.1 (0.97)		4.3 (1.1)	
Catalano, 2003 [16]	Controls: 220 (119 male) IDM: 195 (all GDM) (100 male)	<72 hours	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
			4.1 (1.1)	4.2 (0.8)	4.8 (1.1)	4.6 (1.1)	4.4 (1.2)	4.7 (1.2)	5.5 (1.4)	5.3 (1.4)
			Pooled		**Pooled**		**Pooled**		**Pooled**	
			4.2 (1.0)		4.7 (1.1)		4.6 (1.2)		5.4 (1.4)	
Clarson, 1989 [58]	Controls: 11 IDM: 30 (11 GDM, 19 pre-existing)	<48 hours	Data for sum of 8 skinfolds (midbiceps, midtriceps, subscapular and suprailiac on both sides) Control: 29.0 (5.0) IDM: 33.6 (7.3)							
Enzi, 1980 [25]	Controls: 17 IDM: 25 (8 T1D, 17 GDM)	Birth	Data for sum of 4 skinfolds (subcostal, subscapular, triceps, crural) Control: 17.8 (2.9), IDM: 25.5 (6.6)							
Greco, 2003 [32]	Controls: 16 ODM: 15 (GDM, T1D, T2D)	<24 hours	4.1 (0.1)		4.7 (0.1)		3.8 (0.1)		4.7 (0.1)	
Hollingsworth, 1991 [33]	Controls: 211 ODM: 70 (all GDM)	Newborn - time not stated	3.8 (0.3)		4.3 (0.9)		4.0 (0.4)		5.0 (1.1)	
Krishnaveni, 2005 [30]	Controls: 545 (266 male) IDM: 41 (all GDM) (15 male)	<72 hours	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
			4.1 (0.9)	4.2 (0.9)	4.8 (0.8)	5.2 (1.2)	4.4 (0.9)	4.5 (1.0)	5.1 (1.0)	5.4 (1.1)
			Pooled		**Pooled**		**Pooled**		**Pooled**	
			4.2 (0.9)		5.1 (1.1)		4.4 (0.9)		5.3 (1.1)	

McFarland, 1998 [42]	Controls: 58 (40 male) IDM: 12 (12 GDM, 4 pre-existing) (8 male)	<24 hours	5.4 (1.2)		7.5 (1.5)		5.4 (1.3)		6.8 (2.0)			
Metzger (HAPO), 2009 [15]	Controls: 16097 IDM: 3082 (all GDM)	<72 hours	4.08 (0.86)		4.42 (0.99)		4.17 (0.95)		4.59 (1.13)			
Mohamed, 2010 [59]	Controls: 20 (10 male) IDM: 40 (all T1D) (26 male)	Birth	1.3		2.32 (0.49)		0.48		1.43 (0.55)			
Nasrat, 1997 [60]	Controls: 501 (236 male) IDM: 50 (40 GDM, 10 T2D) (29 male)	<12 hours	3.12		3.26		3.4		3.8			
Nelson, 2007 [34]	Controls: 19 IDM: 56 (all T1D)	Birth	6.0 (2.4)		8.0 (3.1)		5.6 (2.0)		7.4 (2.1)			
Ng, 2004 [61]	Controls: 40 IDM: 80 (68 GDM, 12 T1D)	Birth	TR and SS presented as median and IQR in control, diet treated and insulin treated diabetes groups. No statistically significant differences between the groups.									
Petersen, 1990 [62]	Controls: 16 IDM: 20 (5 GDM, 15 pre-existing diabetes)	<48 hours	Data for sum of 5 skinfolds only; quadriceps, pectoralis, biceps, subscapular and triceps Control: 21.2 (3.6), IDM: 26.8 (5.2)									
Rossi, 2000 [35]	Control: 13 IDM: 6 (GDM and T1D)	Birth	4.76 (0.28)		5.78 (1.5)		4.47 (2.2)		5.78 (1.25)			
Simmons, 1995 [36]	Control: 125 (61 male) IDM: 35 (34 GDM, 1 T2D) (14 male)	<24 hours	4.2 (0.8)		4.5 (1.0)		5.3 (1.3)		6.3 (2.0)			
Sletner, 2013 [27]	Control: 457 (230 male) IDM: 65 (all GDM) (32 male)	<72 hours	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**		
			4.36 (0.99)	4.44 (1.01)	4.70 (1.21)	4.72 (1.05)	4.19 (0.90)	4.34 (1.11)	4.78 (1.33)	4.90 (1.33)		
			Pooled		**Pooled**		**Pooled**		**Pooled**			
			4.40 (1.00)		4.71 (1.12)		4.26 (1.02)		4.84 (1.32)			
Stevenson, 1991 [37]	Control: 20 IDM: 13 (all GDM)	<72 hours	4.3 (0.8)		5.0 (1.1)		No SS measurement taken					
Verma, 1991 [63]	750 newborns (males: 413) No control/ODM numbers	<48 hours	3.73 (0.88)		3.92 (0.83)		3.75 (0.83)		3.99 (0.85)			
Vohr, 1995 [38]	Controls: 143 (76 male) IDM: 119 (all GDM) (59 male)	Birth	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**		
			3.9 (1.0)	4.2 (1.2)	4.0 (1.0)	4.0 (1.2)	4.5 (1.3)	4.8 (1.3)	4.6 (1.4)	4.8 (1.5)		
			Pooled		**Pooled**		**Pooled**		**Pooled**			
			4.1 (1.0)		4.1 (1.1)		4.6 (1.3)		4.7 (1.5)			

Westgate, 2006 [39]	Controls: 61 IDM: 136 (108 GDM, 28 T2D)	<24 hours	4.4 (1.0)		5.1 (1.2)		4.4 (1.0)		5.6 (1.7)			
Whitelaw, 1977 [64]	Controls: 45 (AGA infants) IDM: 18	<48 hours 1 year	Data for sum of 8 skinfolds (biceps, triceps, subscapular and suprailiac on both sides) Birth: Control: 28.3 (4.2) IDM: 41.3 (10.9) 1 year: Control: 26.9 IDM: 24.0									
Wurster, 1984 [40]	Control: 10 (7 male) IDM: 10 (GDM and pre-existing) (3 male)	<48 hours 2 weeks 1 month	5.0 (1.5) 5.0 (1.0) 5.8 (1.2)		5.9 (1.5) 6.2 (1.1) 6.3 (0.8)		5.2 (1.6) 6.0 (1.1) 6.6 (1.2)		6.6 (1.9) 6.6 (1.2) 6.9 (1.1)			
Zhao, 2013 [21]	Controls: 284 (139 male) IDM: 160 (all GDM) (90 male)	<48 hours	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**		
			4.1 (0.1)	4.3 (0.2)	5.1 (0.2)	5.3 (0.3)	3.9 (0.1)	4.0 (0.2)	5.1 (0.3)	5.2 (0.2)		
			Pooled		**Pooled**		**Pooled**		**Pooled**			
			4.4 (0.3)		5.2 (0.4)		4.0 (0.2)		5.2 (0.3)			

Supplementary, table 1: Triceps and subscapular skinfold thickness in infants of mothers with and without diabetes from individual studies included in the systematic review

Author, year, reference	Study details	Diabetes definition and treatment	Body composition measurement	Details of adjusted analyses
Aman, 2011 [31]	PC; single centre, Sweden; GDM and T1D. Control: term infants of mothers with negative screening for GDM in pregnancy matched for gestational age and mode of delivery. Blinding not stated. Newcastle-Ottawa score 4.	GDM defined as 2h glucose >11 mmol/l on 75g OGTT. All GDM mothers treated with diet and multiple pre-meal insulin injections. BG self-monitored daily and mothers monitored twice monthly in clinic with treatment adjusted to blood glucose.	Examined by single study nurse. Harpenden caliper used for skinfold thickness. Body fat mass calculated from birth weight, birth length and flank skinfold thickness using equation proposed by Catelano.	No
Au, 2013 [18] (additional data provided by authors)	Cross-sectional; single centre, Sydney, Australia; GDM. Control: singleton, term infants with no congenital anomalies whose mothers had normal glucose tolerance. Exclusions: pre-existing diabetes. Babies admitted to NICU > 2 days excluded. Blinding not stated. Newcastle-Ottawa score 3.	Diagnosis of GDM based on the Australasian Diabetes In Pregnancy Society (ADIPS) criteria at time of study. Dietary and physical activity advice given and BG monitored four times daily. 3rd trimester pre-and post-prandial mean BGs obtained or progress notes of BGs reviewed. HbA_{1c} levels obtained. Insulin commenced when glycemic targets not met on dietary adjustment.	Neonatal body fat %, fat mass, and fat-free mass assessed at birth using air displacement plethysmography. Adjusted for infancy (refs 4, 6, 7 in text).	Associations between GDM status and body composition investigated using linear regression unadjusted and adjusted for potential confounders: gestational age, infant sex, maternal age, pre-pregnancy BMI, gestational weight gain, parity, smoking, ethnicity, and hypertension.
Brans, 1983 [57]	PC; single centre, USA: GDM, T1D and T2D. 22 of 61 infants of mothers with diabetes were macrosomic (birth weight >90th percentile).	GDM diagnosed by 3hr 100g OGTT in 3rd trimester (at least 2 BGs exceeding thresholds; fasting 105 mg/dL, 1h 195	Skinfold thickness measured on the left side using a Harpenden caliper by the same investigator. Sum of measurements to the	No

	Control: infants of mothers with no diabetes (antenatal fasting BG <100 mg/dL) and no family history of diabetes. 28 of 52 control infants were macrosomic. Blinding not stated. Newcastle-Ottawa score 2.	mg/dL, 2h 165 mg/dL, 3h 145 mg/dL). All GDM women treated with diet and 23 treated with insulin. All described as good control and only 10 had raised HbA1C.	nearest 0.05mm presented.	
Brumbaugh, 2013 [19]	PC; multi-centre, USA; GDM (all had BMI >30 kg/m2) Control: term infants of mothers with negative OGTT at 24-28 weeks of pregnancy and BMI <25 kg/m^2. Exclusions: pre-existing diabetes and IUGR infants. Blinding not stated. Newcastle-Ottawa score 2.	GDM diagnosed by 3hr OGTT using Carpenter and Coustan criteria. Treatment not described.	Triceps and subscapular skin fold thickness obtained in triplicate by single investigator with a Lange caliper. Total fat mass and fat-free mass measured with air displacement plethysmography. Adjusted for infancy (refs 22, 23 in text). MRI for subcutaneous and intra-abdominal fat.	No
Brunner, 2013 [28] (additional data provided by authors)	Munich population study, Germany; GDM. Data from a follow up of an RCT of LCPUFA supplementation in pregnancy (controls only presented); Control: singleton term infants of mothers with normal glucose tolerance and pre-pregnancy BMI 18-30. Exclusions: pre-existing diabetes. Blinding to maternal diabetes status not described (but not the aim of the study). Newcastle-Ottawa score 3.	For meta-analysis the authors provided the data by maternal diabetes status. GDM diagnosis established from data provided by the women´s gynaecologists. GDM screening was not routine at the time of the study and not all women underwent screening.	Skinfold thickness was measured by trained research assistants in triplicate using a Holtain caliper on the left side and the mean measurement used. Body fat % was calculated using Weststrate's method.	No
Buhling, 2012 [29]	PC; single centre, Germany; GDM. Control: infants of mothers with negative screening or no risk factors for GDM. Blinding not stated. Newcastle-Ottawa score 3.	GDM diagnosed by OGTT and ADA criteria. Treatment not described.	Skinfold thicknesses measured using a single observer with a single measurement and a Holtain caliper.	No adjustments of skin fold thickness results.

Catalano, 2003 [16] (additional data provided by authors)	PC; single centre, USA; GDM. Control: singleton term infants of mothers with 1h 50g glucose screen <135mg/dl or normal 3h 100g OGTT. Exclusions: infants <2000g. Blinding not stated. Newcastle-Ottawa score 2.	GDM diagnosed by National Diabetes Data Group criteria. Diet controlled unless glycaemic target not achieved (fasting glucose <100 mg/dl or 2h postprandial glucose <120mg/dL), then insulin controlled. 67 women with GDM required insulin.	Total body electrical conductivity to assess fat free mass, fat mass and body fat %. 10 measurements taken and averaged. Adjusted for infancy (ref 12 in text). Skinfold thickness assessed using a Harpenden caliper. Anthropometric and TOBEC measurements performed by 1 of 3 examiners.	Unadjusted data presented and adjustments made for gestational age, maternal pregravid weight, weight at last antenatal visit, race, smoking status, and maternal and paternal height.
Clarson, 1989 [58]	PC; single centre, Canada; GDM and pre-existing diabetes. Controls: term infants of mothers with normal glucose tolerance. Blinding not stated. Newcastle-Ottawa score 1.	GDM diagnosed by OGTT at the end of the 2nd trimester using O'Sullivan criteria. Insulin treatment given to all women with GDM from time of diagnosis.	Skinfold thickness measured using a Lange caliper.	No
Durnwald, 2004 [44]	PC; single centre, USA; LGA infants (>90th percentile for gestational age, race and sex) of GDM and non-diabetic mothers. Control: LGA infants of mothers with either normal 1hr 50g OGTT or 3h 100g OGTT. Blinding not stated. Newcastle-Ottawa score 3.	GDM diagnosed by Carpenter and Coustan criteria. Diet controlled unless glycaemic target not achieved (fasting glucose <95 mg/dl or 2h postprandial glucose <120mg/dL), then insulin controlled. 26 of 50 mothers received insulin.	Total body electrical conductivity to assess fat free mass, fat mass and body fat %. 10 measurements taken and averaged. Adjusted for infancy (ref 14 in text). Skinfold thickness assessed using a Harpenden caliper. 92 body composition measurements by TOBEC and 10 by skinfolds.	Unadjusted data presented and stepwise regression analysis used with adjustment for maternal age, pre-pregnancy weight, gestational age and smoking history.

Enzi, 1980 [25]	PC; single centre, Italy; GDM and T1D. Control: term infants of mothers with negative screening for GDM in pregnancy (100g OGTT). Blinding not stated. Newcastle-Ottawa score 1.	Pre-existing diagnosis of T1D or GDM diagnosed by 100g OGTT and National Data Group Criteria. Dietary advice provided and insulin given in T1D group 3 times daily. T1D mothers also hospitalised 6 times during pregnancy for BG assessment.	Skinfold thickness assessed using a Harpenden caliper, and body fat mass calculated from anthropometric measurements by Dauncey's method (infant specific).	No
Greco, 2003 [32]	PC; single centre, Italy; GDM or pre-existing well-controlled (HbA1c 3.5-5.3%) diabetes. Controls: singleton infants of primiparous Caucasian women with no relevant medical history and negative 28 week 100g OGTT, matched for age and pre-pregnancy weight. Blinding not stated. Newcastle-Ottawa score 2.	Diabetes classified using White Class. Intensive monitoring to achieve strict glucose control. Maternal pre-prandial and post-prandial BG assessment 3 times daily. HbA1c monitored every 4 weeks and insulin therapy altered on basis of BG monitoring and HbA1c.	Skinfold measurements taken from the left hand side by a single observer using calipers and the mean of 3 measurements calculated.	No
Hammami 2001 [43]	PC; single centre, USA; GDM or pre-existing diabetes. LGA (≥4000g and >90th percentile) v AGA. IDM in LGA group only. LGA babies presented. Controls: infants of mothers with normal blood glucose screening or normal glucose tolerance test. Blinding not stated. Newcastle-Ottawa score 2.	Mothers with diabetes defined by National Diabetes Data Group criteria. 9 with GDM (7 on insulin), 1 with T2D and 1 with T1D (both on insulin).	DXA whole body scans performed with a single beam densitometer. Adjusted for infancy (ref 7 in text).	Unadjusted data presented. Regression analysis used to investigate weight, length, race and sex as influences of body composition in AGA infants.
Hollingsworth, 1991 [33]	PC; single centre, USA; GDM. Control: infants of mothers with negative 1h 50g glucose screen in pregnancy. Newcastle-Ottawa score 3.	GDM diagnosed by ≥2 abnormal values on 3h 100g OGTT using O'Sullivan criteria. Treatment not stated.	Skinfold thickness measured. No further detail given.	No

Krishnaveni, 2005 [30] (additional data provided by authors)	RC; single centre, India; GDM. Controls: singleton infants of mothers with negative GDM screening. Exclusions: pre-existing diabetes and SGA infants. Blinding not stated. Newcastle-Ottawa score 4.	GDM diagnosed following 100g 3h OGTT using Carpenter and Coustan criteria. All mothers treated with diet and 12 (29%) received insulin.	Skinfold thickness measured by one of four trained observers using Harpenden caliper.	No
Lee, 2012 [20] (additional data provided by authors)	PC; multi-centre, USA; singleton infants of uncomplicated pregnancies for development of adiposity reference ranges. Study included an additional cohort of 25 infants of mothers with GDM or pre-existing diabetes. Exclusions: pregnancies with poor menstrual dating and infants with congenital anomalies. Blinding not stated. Newcastle-Ottawa score 2.	The authors state "An important study limitation was the unavailability of precise diagnostic criteria for classifying diabetic pregnancies." T1D, T2D and GDM included. No treatment information discussed.	Percentage body fat and total body fat mass estimated using direct measurements of volume and body mass based on principles of whole body densitometry. Adjusted for infancy (ref 29 in text).	Unadjusted data supplied by authors. Data adjusted for newborn weight, gestational age, gender and ethnicity using multiple regression. Variables entered the model if p <0.05 and were removed if p >0.10.
Lingwood, 2011 [17] (additional data provided by authors)	PC; single centre, Australia; GDM. Control: singleton term infants of mothers with normal glucose tolerance and BMI (18.5-25) at booking. Recruited in the same centre over a similar time period and assessed by the same methods but reported separately. Exclusions: mothers with a history of maternal illness other than GDM and infants with congenital anomalies. Newcastle-Ottawa score 2.	GDM diagnosis by 75g GTT (fasting ≥5.5 or 2h ≥8.0 mmol/L) and ADIPS criteria. Treated following ADIPS guidelines. Diet and activity advice. BG levels monitored 4x per day with targets of fasting ≤5.5 or 2h ≤8.0 mmol/L. Insulin commenced if necessary. Relatively well-controlled group of women with GDM (80% met both ADIPS targets above on average).	Body composition measured by air displacement plethysmography. Infant body fat % was computed from body density by software integral to the system. Age and sex-specific densities of fat-free mass were computed based on data of Fomon	No

Metzger (HAPO), 2009 [15] (additional data provided by authors)	PC; multi-centre, multinational study, ethnically diverse cohort involving 15 field centres; GDM. Medical caregivers blinded to maternal glucose levels unless following criteria met: fasting plasma glucose >5.8 mmol/L, 2hr OGTT >11.1 mmol/L or random glucose ≥8.9mmol/L. Singleton infants. For purpose of meta-analysis authors grouped GDM and controls using more recent International Association of Diabetes in Pregnancy Study Groups (IADPSG) thresholds (FPG >5.1 mmol/L, 2hr OGTT >8.5 mmol/L.) Blinding occurred as GDM diagnosed post Hoc. Newcastle-Ottawa score 5.	GDM defined retrospectively using IADPSG thresholds. As women with GDM had BG levels below the pre-defined study diabetes thresholds, caregivers remained blinded to their values and no monitoring or treatment was provided.	Rigorous training and annual recertification in anthropometric measurements was established for research nurses and midwives. 2 skinfold measurements were taken using a Harpenden caliper and if results differed by >0.5mm, a 3rd was taken. Averages used. Fat mass calculated using equation by Catalano (infant specific). (4)	Authors provided unadjusted and adjusted data. Adjustments made for field centre, gestational age, maternal parity, neonatal gender and maternal BMI at the time of the OGTT.
McFarland, 1998 [42]	PC; single centre, USA; GDM and pre-existing diabetes. Macrosomic (>4000g) infants only. Control: macrosomic infants of mothers with negative glucose screening during pregnancy. Blinding not stated. Newcastle-Ottawa score 2.	GDM diagnosed by National Diabetes Data Group criteria. Diabetes treated by local protocol (ref 11 in text). BG monitored 7x/day. Glucose control in group demonstrated to be poor (mean BG 125±34 mg/dL).	All skinfold measurements performed by one of two examiners. Estimation of body fat mass performed by Dauncey's method (infant specific).	No
Mohamed, 2010 [59]	PC; single centre, Egypt; T1D. Divided into macrosomic and non-macrosomic groups (> or < 4000g) Control: term infants, <4000g of non-diabetic mothers, matched for age and sex. Glucose screening in pregnancy not discussed. Exclusions: included infection, congenital or chromosomal abnormalities and metabolic disorders.	TID diagnosed pre-pregnancy. No further details.	Skinfold thickness measured using a skinfold caliper.	No

	Blinding not stated. Newcastle-Ottawa score 2.			
Nasrat, 1997 [60]	PC; single centre, Saudi Arabia; GDM and pre-existing diabetes. Control: singleton infants of mothers with negative glucose screening in pregnancy. Blinding not stated. Newcastle-Ottawa score 3.	GDM diagnosed following 100g 3h OGTT using NGGD criteria. 9 of 40 GDM and 0 of 10 T2D required insulin.	Skinfold thickness measured using a Harpenden caliper by one of 2 trained examiners.	No
Nelson, 2007 [34]	PC; multi-centre study, UK; T1D. Control: infants of mothers with negative routine screening for gestational diabetes (national guidelines: http://www.sign.ac.uk/guidelines/fulltext/55/section8.html). Blinding not stated. Newcastle-Ottawa score 4.	T1D diagnosis and treatment not described.	Skinfold thickness at subscapular and triceps measured using a Holtain caliper by paediatricians using a centrally agreed protocol.	No
Ng, 2004 [61]	PC; single centre, Hong Kong; GDM and T1D. Control: term infants admitted to neonatal unit with suspected perinatal infection but with negative cultures and normal CRP (<10mg/l). GDM screening not discussed. Exclusions: infants with chromosomal or congenital abnormalities. Blinding not stated. Newcastle-Ottawa score 1.	T1D diagnosed prior to study. GDM diagnosed by fasting glucose >5.5mmol/L or >8.0mmol/L 2h after 75g OGTT. GDM treated with low energy diet (1800kcal/d) and T1D treated with daily insulin.	Skinfold thickness measured using a Holtain skinfold caliper.	No

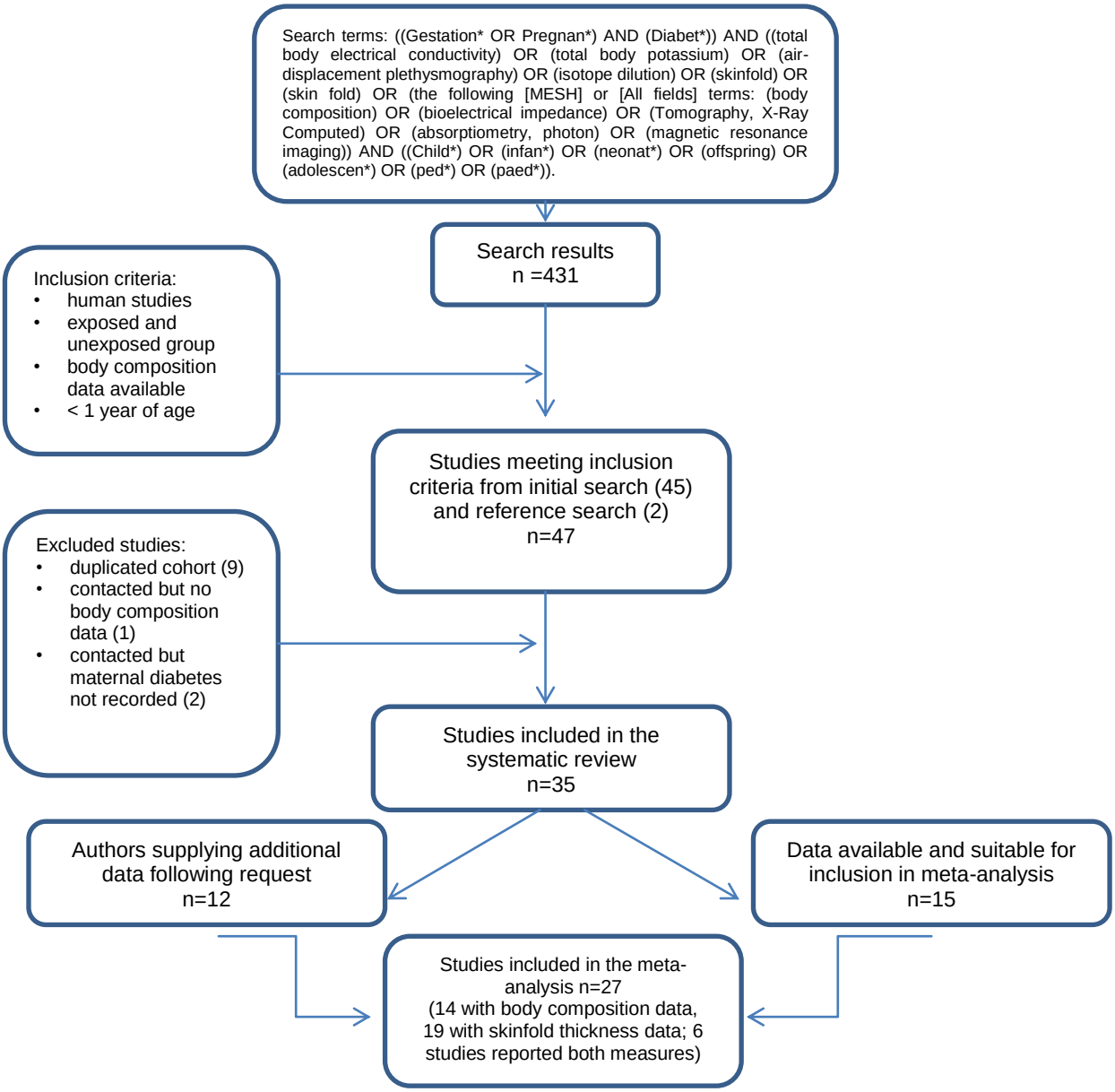
Petersen, 1990 [62]	PC; single centre, Denmark; GDM and pre-existing diabetes. 3 groups of infants: SGA (<10th percentile), infants of mothers with diabetes and controls. Control: term AGA infants of mothers with normal glucose tolerance. Blinding not stated. Newcastle-Ottawa score 1.	Maternal diabetes categorized by White classification. (GDM 5, pre-existing 15) Further details of diagnosis or treatment not given.	Skinfold thickness measured using a Harpenden caliper.	No
Rossi, 2000 [35] (Italian)	PC; single centre, Italy; GDM and T1D. Control: women without diabetes. Blinding not stated. Newcastle-Ottawa score 2.	No information provided on the diagnosis and management of GDM. T1D treated with insulin.	Skinfold thickness measured. No further details provided.	No
Simmons, 1995 [36]	PC; two centres, New Zealand (Maori, Indian and Pacific Islanders); GDM or previous diet treated diabetes. Controls: singleton term infants of non-diabetic mothers with no other medical illness and smoking <10 cigarettes/day. Exclusions: mothers with pre-eclampsia and IUGR infants. European diabetic women excluded as numbers small. Blinding not stated. Newcastle-Ottawa score 3.	GDM diagnosed using O'Sullivan criteria. Women with diabetes maintained on 1800 cal diet. Those with continuing hyperglycaemia (fasting BG >5.5 mmol/L, 2h post prandial BG >6.5 mmol/L) despite therapy were treated with insulin. 15% required insulin.	Skinfold thickness measured using a Holtain paediatric caliper. All measures taken by one observer on the left side and the mean of 3 recordings was taken.	Data adjusted for neonatal sex and maternal ethnic group.
Schaefer-Graf, 2011 [26]	PC; single centre, Germany; GDM from previous study (ref 2 in paper) compared with controls. Controls: singleton infants of mothers with normal 75g OGTT.	GDM diagnosed by 75g GTT using Carpenter and Coustan criteria (2 values above 5.0/10.0/8.6 mmol/l). Women given dietary advice and	Skinfold thickness measured at the flank and fat mass calculated by formula derived by Catalano (infant specific).	Unadjusted body composition data presented.

	Exclusions: infants <34 weeks or with congenital anomalies. Blinding not stated. Newcastle-Ottawa score 3.	performed self-monitoring of BG. Insulin therapy given based on BG levels or glycaemia plus fetal growth (randomised in initial trial). Equivalent outcomes in both groups of GDMs.		
Sletner, 2013 [27] (additional data supplied by authors)	Population-based cohort; multi-centre study, Norway; GDM. Authors separated skinfold results based on maternal GDM status for purpose of review. Controls: infants of mothers with negative 75g screening OGTT at 28 weeks gestation. Exclusions: pre-existing diabetes. Blinding not stated. Newcastle-Ottawa score 4.	GDM cases were diagnosed using WHO criteria. Women with FPG >7.0 mmol/l or 2h BG >9.0 mmol/l were referred to secondary care and those with 2h BG 7.8–9.0 mmol/l were referred to their GP after lifestyle advice had been given. Authors also provided results by IADPSG criteria.	Measurements performed by specially trained personnel. Skinfolds measured using a Holtain caliper. Measurements performed twice to nearest 0.1 mm and means used.	Unadjusted data presented. Authors also provided data adjusted for maternal BMI.
Stevenson, 1991 [37]	PC; single centre, USA; GDM. AGA and LGA infants recruited. Controls: term, AGA age infants. Blinding not stated. Newcastle-Ottawa score 1.	GDM diagnosed by glucose tolerance test using O'Sullivan and Mahan criteria. Dietary control then initiated.	Triceps skin fold measured by the same individual using a Lange caliper with mean taken of 3 separate measurements.	No
Verma, 1991 [63]	PC; single centre; India. All infants delivered consecutively included. Preterms and SGA included. Maternal antenatal history including diabetes recorded. Blinding not stated. Newcastle-Ottawa score 3.	Maternal complications in the antenatal period, including diabetes recorded. No further details and numbers of infants of mothers with diabetes not given.	Skinfold thickness using a Harpenden caliper.	No

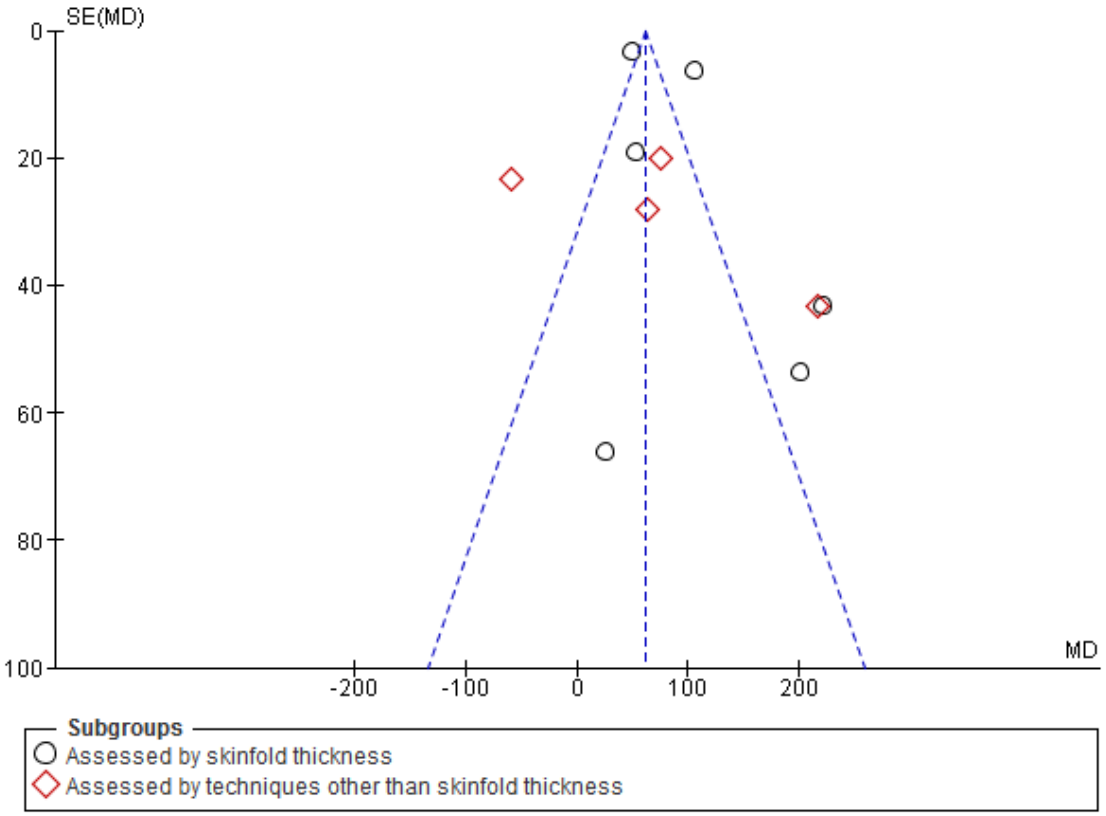
Vohr, 1995 [38] (additional data provided by authors)	PC; single centre, USA; GDM enrolled to include equal numbers of LGA and AGA babies (separated into LGA and AGA groups). Controls: infants of mothers with negative 1h 50g glucose screen at 24-28 weeks, enrolled to include equal numbers of LGA and AGA babies. Exclusions: infants with congenital anomalies or requiring intensive care. Blinding not stated. Newcastle-Ottawa score 3.	GDM diagnosed by 100g OGTT and Carpenter and Coustan criteria. Initial dietary advice and weekly BG monitoring aiming to keep fasting glucose <100mg/dl and all other values <120mg/dl. Insulin therapy recommended if these levels were exceeded.	Skinfolds measured by 1 of 2 trained research nurses, mean taken of 2 measurements using a Lange caliper.	No
Vohr, 1997 [41] (additional data provided by authors)	PC (subgroup of Vohr 1995); single centre, USA; GDM enrolled to include equal numbers of LGA and AGA babies (separated into LGA and AGA groups). Controls: infants of mothers with negative 1h 50g glucose screen at 24-28 weeks, enrolled to include equal numbers of LGA and AGA babies. Exclusions: infants with congenital anomalies or requiring intensive care. Blinding not stated. Newcastle-Ottawa score 2.	GDM diagnosed by 100g OGTT and Carpenter and Coustan criteria. Initial dietary advice and weekly BG monitoring aiming to keep fasting glucose <100mg/dl and all other values <120mg/dl. Insulin therapy recommended if these levels were exceeded.	Skinfolds measured by 1 of 2 research nurses, mean taken of 2 measurements using Lange caliper.	No
Westgate, 2006 [39]	PC; single centre, New Zealand (Maori and Pacific island mothers); GDM and T2D. Control: infants of mothers with 1h 50g glucose screen <7.8mmol/L. Mothers invited using pseudo-random number generator against clinic lists and matched by ethnicity. Blinding not stated.	T2D diagnosed prior to pregnancy. GDM diagnosed by fasting glucose ≥5.5mmol/L or 2h value ≥9.0. BG monitoring at home (5 tests daily), with insulin commenced if fasting BG >5.5 mmol/L on 2 occasions or post-prandial readings consistently ≥6.5 mmol/L.	Neonatal paediatrician or senior neonatal nurse assessed babies using a Harpenden caliper for skinfold thickness.	No

	Newcastle-Ottawa score 3.	Insulin pump used if daily dosage >200 units.		
Whitelaw, 1977 [64]	PC; single centre, UK; 5 groups of infants studied: SGA (<10th percentile), AGA (10th-90th percentile), LGA (>90th percentile), infants of mothers with diabetes and infants of obese mothers. Control: AGA infants of mothers with normal glucose tolerance. Newcastle-Ottawa score 2.	No details of diabetes type, diagnosis, or treatment.	Skinfolds measurements by a single observer using a Harpenden caliper.	No
Wurster, 1984 [40]	PC; single centre, USA; GDM or pre-existing diabetes. Controls: infants of mothers without diabetes. No blinding stated. Newcastle-Ottawa score 2.	No details of diabetes diagnosis. Blood glucose monitored 4 times a day during pregnancy with an aim of maintaining pre-meal BG <100 mg/dL. Achieved in majority and women described as well-controlled diabetics.	Skinfold measurements taken on the left side of the body using a Harpenden caliper.	Unadjusted data presented and further subgroup analysis performed to confirm that infant sex was not an influence.
Zhao, 2013 [21] (Chinese) (additional data provided by authors)	PC; single centre, China; GDM v control in AGA (birth weight 10-90th percentile of local population for sex and gestation) infants. Blinding not stated. Newcastle-Ottawa score 3.	GDM diagnosis based on 1998 WHO diagnostic definition.	Anthropometric and skinfold measurements obtained and fat mass and fat-free mass calculated.	Unadjusted data presented and regressions performed of factors associated with differences in the anthropometric measurements but adjusted results not provided.

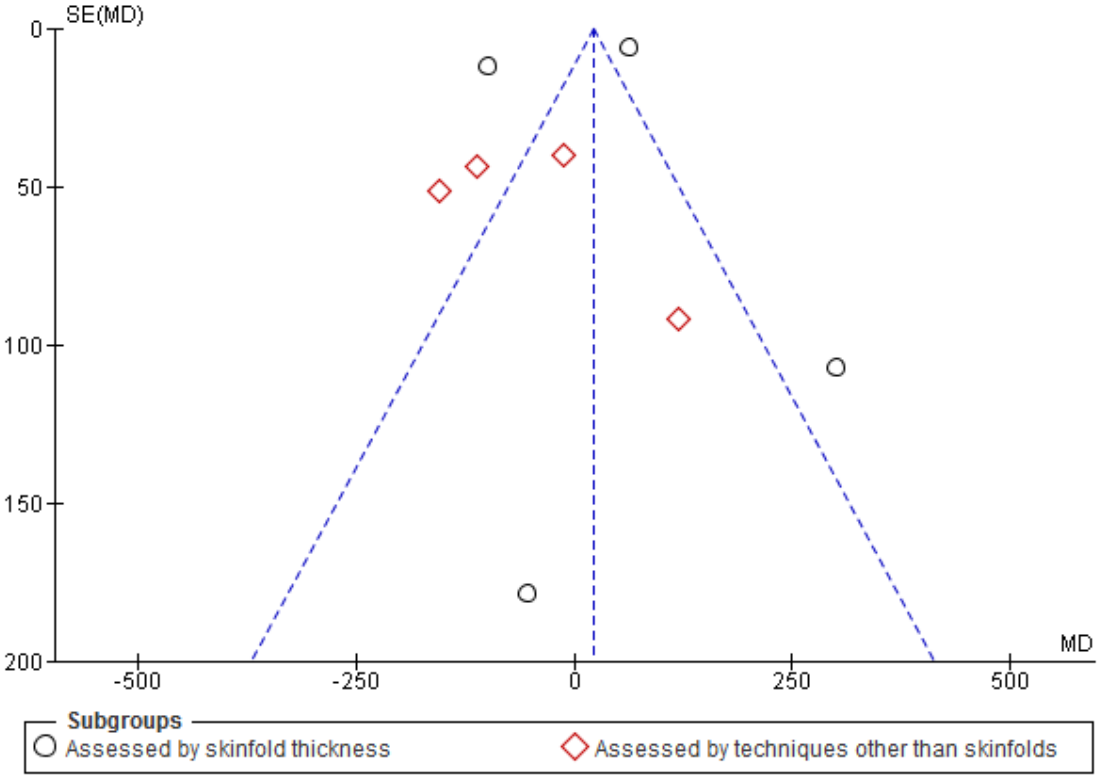
Supplementary, table 2: Studies included in the systematic review examining the association between intrauterine exposure to maternal diabetes and offspring adiposity in infants. GDM: gestational diabetes mellitus, T1D: type 1 diabetes, T2D: type 2 diabetes.



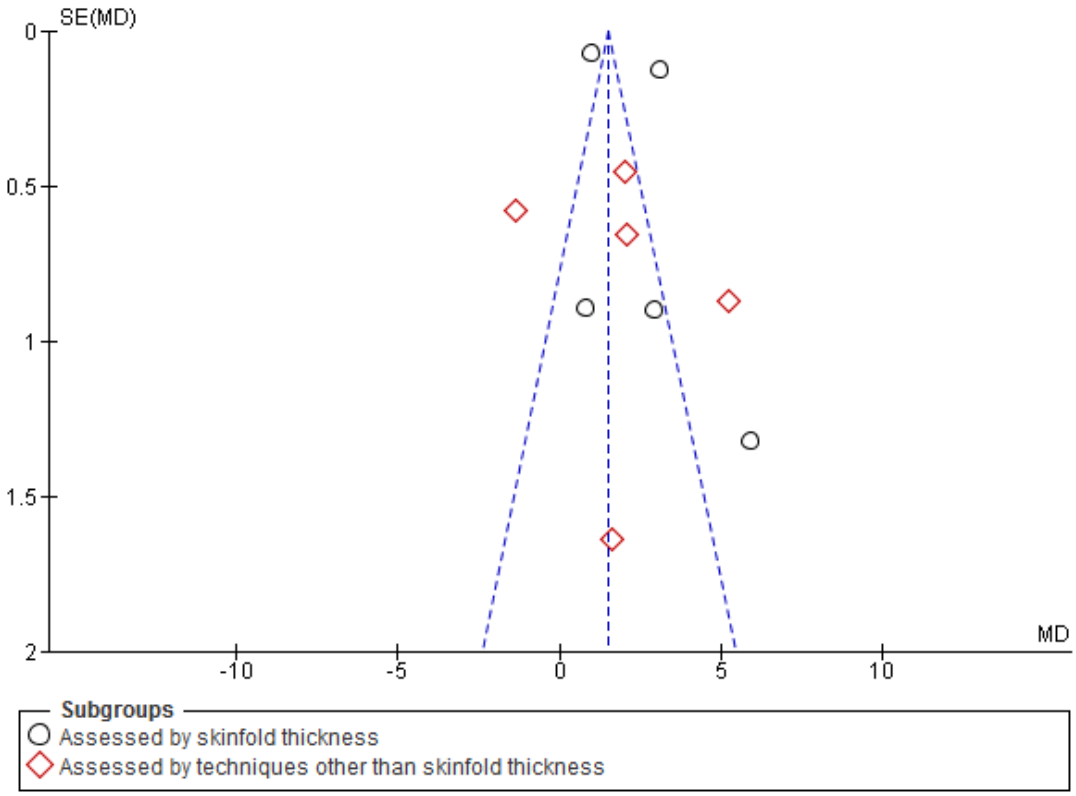
Supplementary figure 1 Flow chart of the search strategy used in this review. The relevant number of papers at each point is given.



Supplementary figure 2 Funnel plot of studies comparing fat mass (g) in IDM and NIDM (all diabetes types)



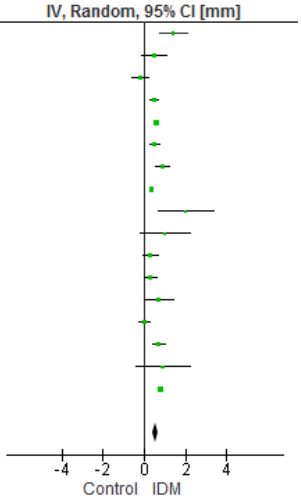
Supplementary figure 3 Funnel plot of studies comparing fat-free mass (g) in IDM and NIDM (all diabetes types)



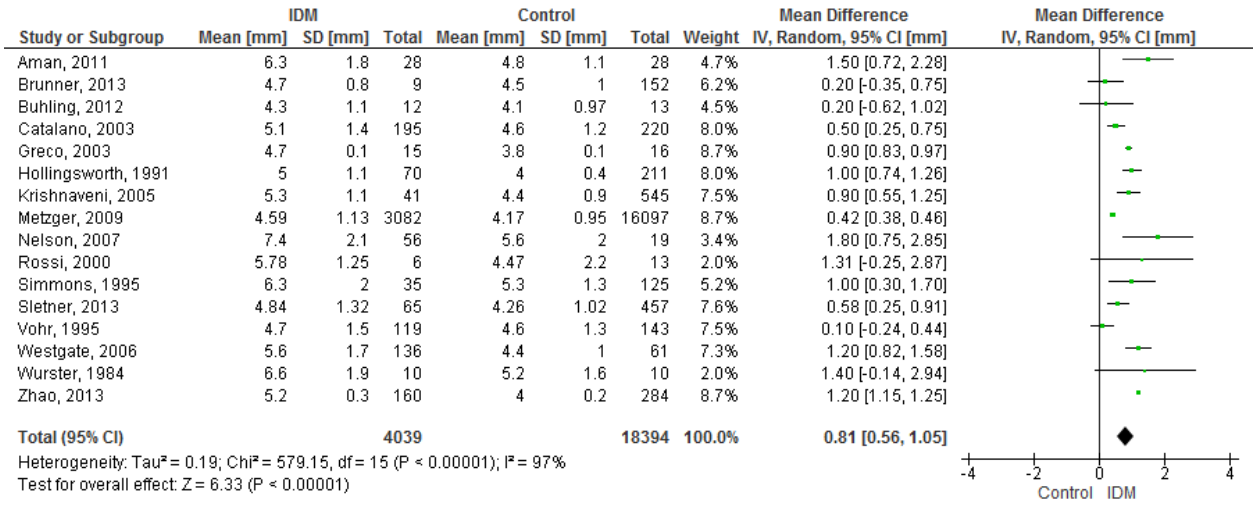
Supplementary figure 4 Funnel plot of studies comparing body fat % in IDM and NIDM (all diabetes types)

Study or Subgroup	IDM Mean [mm]	IDM SD [mm]	IDM Total	Control Mean [mm]	Control SD [mm]	Control Total	Weight	Mean Difference IV, Random, 95% CI [mm]
Aman, 2011	6.7	1.4	28	5.3	1.1	28	3.4%	1.40 [0.74, 2.06]
Brunner, 2013	5.2	0.9	9	4.7	0.9	152	3.8%	0.50 [-0.11, 1.11]
Buhling, 2012	4.6	0.9	30	4.8	1.5	142	5.8%	-0.20 [-0.61, 0.21]
Catalano, 2003	4.7	1.1	195	4.2	1	220	8.4%	0.50 [0.30, 0.70]
Greco, 2003	4.7	0.1	15	4.1	0.1	16	9.6%	0.60 [0.53, 0.67]
Hollingsworth, 1991	4.3	0.9	70	3.8	0.3	211	8.2%	0.50 [0.29, 0.71]
Krishnaveni, 2005	5.1	1.1	41	4.2	0.9	545	6.5%	0.90 [0.55, 1.25]
Metzger, 2009	4.42	0.99	3082	4.08	0.86	16097	9.8%	0.34 [0.30, 0.38]
Nelson, 2007	8	3.1	56	6	2.4	19	1.1%	2.00 [0.65, 3.35]
Rossi, 2000	5.78	1.5	6	4.76	0.28	13	1.3%	1.02 [-0.19, 2.23]
Simmons, 1995	4.5	1	35	4.2	0.8	125	6.3%	0.30 [-0.06, 0.66]
Sletner, 2013	4.71	1.12	65	4.4	1	457	7.3%	0.31 [0.02, 0.60]
Stevenson, 1991	5	1.1	13	4.3	0.8	20	3.2%	0.70 [0.01, 1.39]
Vohr, 1995	4.1	1.1	119	4.1	1.1	143	7.5%	0.00 [-0.27, 0.27]
Westgate, 2006	5.1	1.2	136	4.4	1	61	6.8%	0.70 [0.38, 1.02]
Wurster, 1984	5.9	1.5	10	5	1.5	10	1.2%	0.90 [-0.41, 2.21]
Zhao, 2013	5.2	0.4	160	4.4	0.3	284	9.6%	0.80 [0.73, 0.87]
Total (95% CI)			**4070**			**18543**	**100.0%**	**0.52 [0.37, 0.68]**

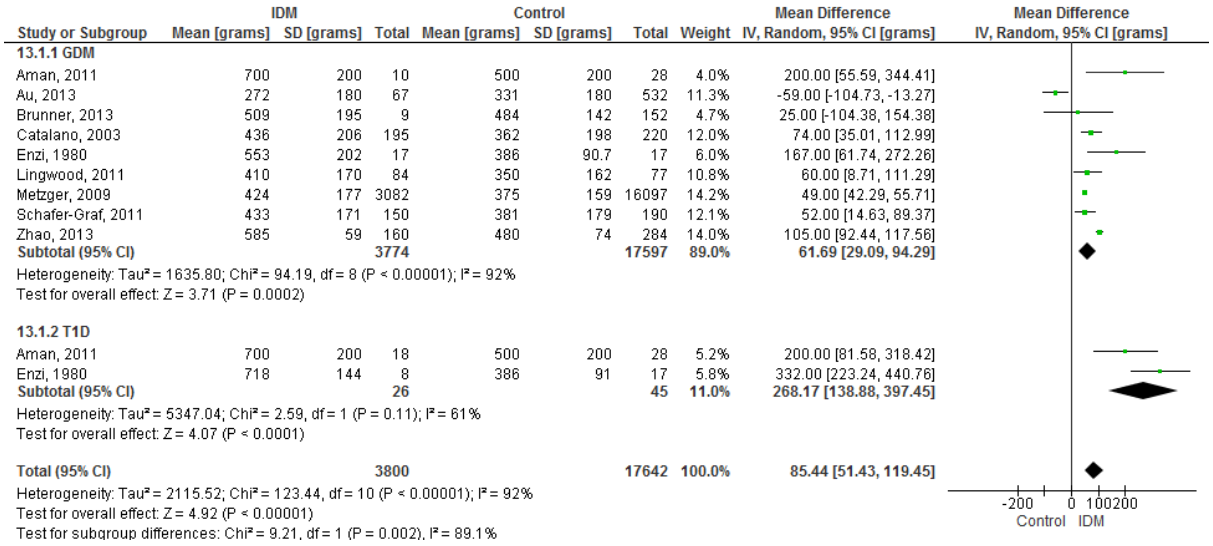
Heterogeneity: Tau² = 0.06; Chi² = 188.84, df = 16 (P < 0.00001); I² = 92%
Test for overall effect: Z = 6.79 (P < 0.00001)



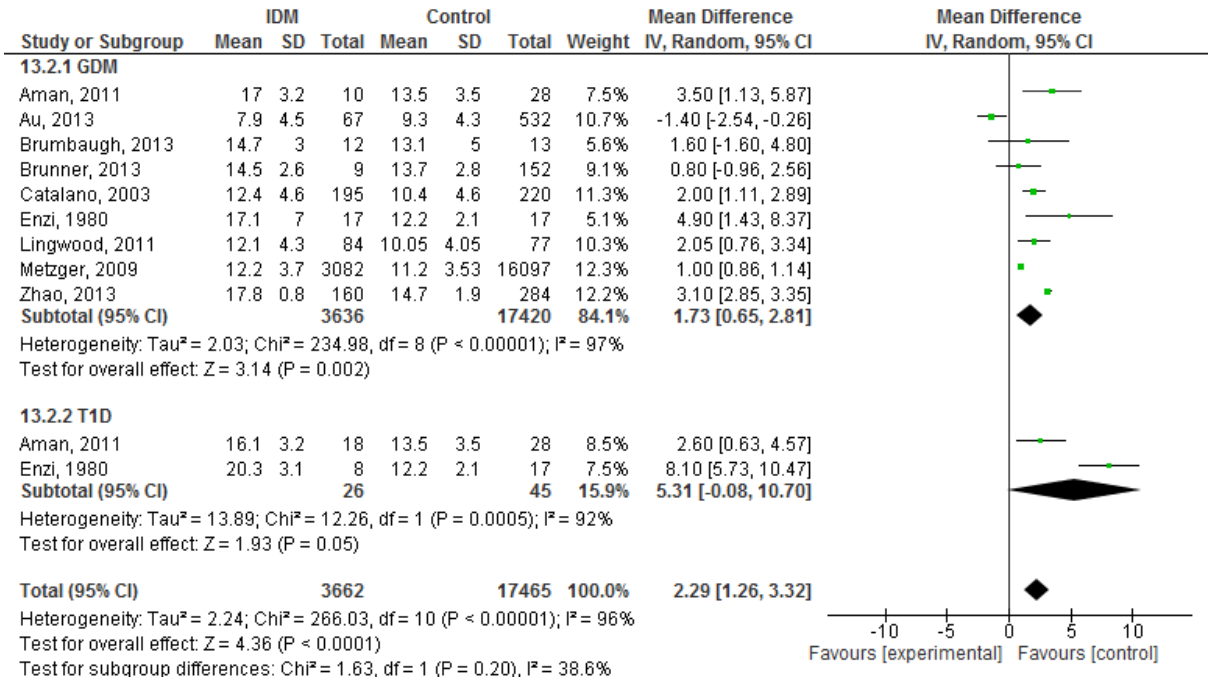
Supplementary figure 5 Forest plot (random effects analysis) comparing triceps skinfold thickness (mm) in IDM and NIDM (all types of diabetes)



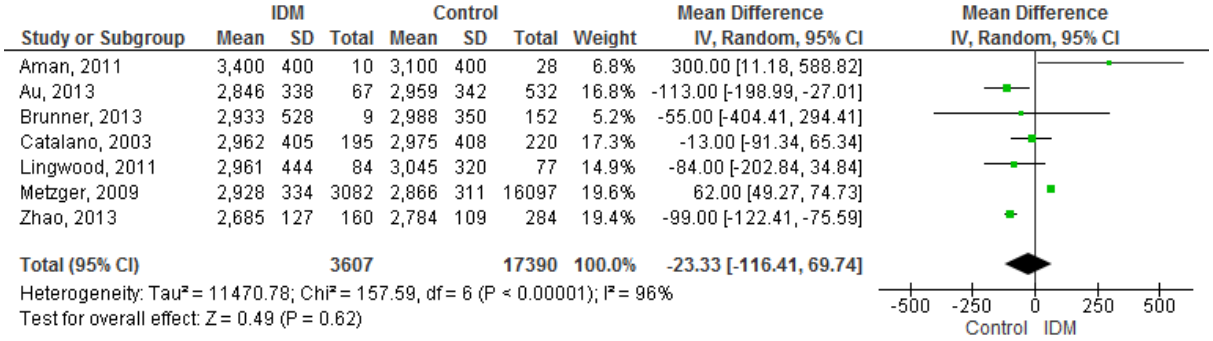
Supplementary figure 6 Forest plot (random effects analysis) comparing subscapular skinfold thickness (mm) in IDM and NIDM (all types of diabetes)



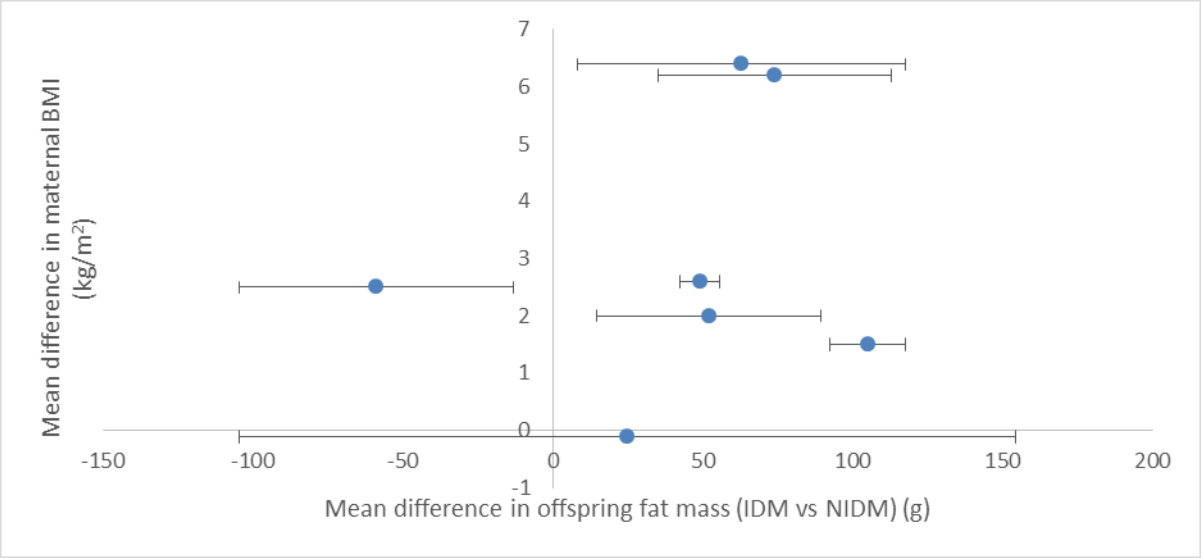
Supplementary figure 7 Forest plot (random effects analysis) comparing fat mass (g) in infants of mothers with and without diabetes with subgroup analysis by maternal diabetes type (GDM and T1D)



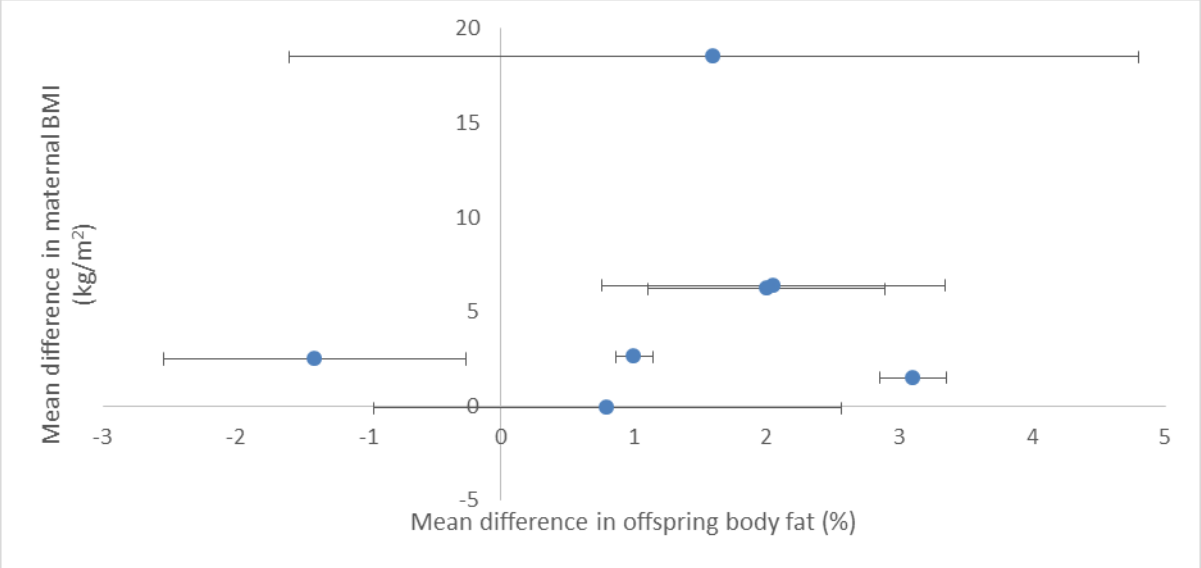
Supplementary figure 8 Forest plot (random effects analysis) comparing body fat % in infants of mothers with and without diabetes with subgroup analysis by maternal diabetes type (GDM and T1D)



Supplementary figure 9 Forest plot (random effects analysis) comparing fat-free mass (g) in infants of mothers with and without GDM



Supplementary figure 10 Mean difference in maternal BMI between mothers with and without diabetes in each study against mean difference in infant fat mass, showing no evidence of increasing fat mass with increasing maternal BMI.



Supplementary figure 11 Mean difference in maternal BMI between mothers with and without diabetes in each study against mean difference in infant body fat %, showing no evidence of increasing body fat % with increasing maternal BMI.