NCC-WCH

Version 1

Diabetes in Pregnancy (update) Appendices – Set 2

Evidence tables

Clinical Guideline <…>

Methods, evidence and recommendations

01 September 2014

Draft for Consultation

Commissioned by the National Institute for Health and Care Excellence

Disclaimer
Healthcare professionals are expected to take NICE clinical guidelines fully into account when exercising their clinical judgement. However, the guidance does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of each patient, in consultation with the patient and/or their guardian or carer.

Funding
<Please insert your statement here

Contents

Appendices

Appendix A: Evidence Tables

A.1 What is the effectiveness of oral oestrogen-containing or progestogen-containing contraceptives in women with diabetes compared with women without diabetes?

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation Ahmed,S.B., Hovind,P., Parving,H.H., Rossing,P., Price,D.A., Laffel,L.M., Lansang,M.C., Stevanovic,R., Fisher,N.D., Hollenberg,N.K., Oral contraceptives, angiotensin-dependent renal vasoconstriction, and risk of diabetic nephropathy, Diabetes Care, 28, 1988-1994, 2005 Ref Id 203342 Country/ies where the study was carried out United States of America Study type Comparative observational study Aim of the study The study aimed to: 1) investigate the renal	Sample size Whole study: n= 92 (Women with diabetes using oral contraceptives= 12 Women without diabetes using oral contraceptives= 10 Women with diabetes not using oral contraceptives= 29 Women without diabetes not using oral contraceptives= 41) Subgroup of interest to the NCC-WCH review: n= 22 (Women with diabetes using oral contraceptives= 10 Women without diabetes using oral contraceptives= 12) Characteristics No participants were taking medication other than oral contraceptives, oral hypoglycemic agents, angiotensin converting enzyme (ACE) inhibitors, or angiotensin receptor blockers (ARBs).	Interventions None* *This study was performed in two parts. The first was a comparative observational study comparing four groups of women: women with diabetes taking oral contraceptives, women without diabetes taking oral contraceptives, women with diabetes not taking oral contraceptives, and women without diabetes not taking oral contraceptives. The second part was an intervention study on the use of captopril in women with diabetes. The intervention, methods and results for the first part of the study are reported here.	Details The methods used in the first part of the study are reported here.* All participants gave written informed consent. Approval for the study protocol was granted by the Brigham and Women's Hospital Institutional Review Board. An initial medical history, physical examination, electrocardiogram, and laboratory screening was performed on all participants. ACE inhibitors and ARBs were discontinued for two weeks prior to the study. Participants consumed >200mmol sodium/day for 4 days prior to the study (no data were excluded due to dietary non-compliance). A 24 hour urine collection was used	Results The results for the first part of the study are reported here.* Results are mean±standard error (SE) unless otherwise stated Results at baseline: Mean arterial pressure (mmHg) Women with diabetes= 83±2 Women without diabetes= 87±2 Fasting plasma glucose (mmol/l) Women with diabetes= 8.33±1.17 (reported as 150±21 mg/dl in the study paper) Women without diabetes= 4.4±0.17 (reported as 79±3 mg/dl in the study paper) p<0.05 HbA_{1c} (%)	Limitations It is not clear how participants were recruited into the study The inclusion and exclusion criteria were not reported NICE guidelines manual. Appendix I: Methodology checklist: Prognostic studies 1.1 The study sample represents the population of interest with regard to key characteristics, sufficient to limit potential bias to the results - Yes 1.2 Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias -

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
plasma flow response to captopril, as an index of renin angiotensin system activity, 2) determine whether the use of oral contraceptives in women newly diagnosed with type 1 diabetes is associated with the development of nephropathy Study dates September 1979 to August 1984 Source of funding One author was supported by a biomedical fellowship from the Kidney Foundation of Canada The study was supported by grants from the National Institutes of Health to one author. The second part of the study was carried out with financial support from the Danish Diabetes Association, the Paul and Erna Sehested Hansen Foundation, the Aase and Ejnar Danielsen Foundation, and the Per S. Henriksen Foundation	Characteristics of all included women in the study (n=92): Age (years) Women with diabetes= 24±2 Women without diabetes= 27±2 Body Mass Index (BMI) Women with diabetes= 26±1.7 Women without diabetes= 29±2.4 Smokers Women with diabetes= 3/12 (25%) Women without diabetes= 1/10 (10%) Oral contraceptive estrogen content (µg/tablet) Women with diabetes= 31.0±1.9 Women without diabetes= 30.5±2.1 Oral contraceptive progesterone content (mg/tablet) Women with diabetes= 0.34±0.11 Women without diabetes= 0.36±0.12 Known duration of diabetes in diabetes group= 9.5 years±1.3		to measure sodium, creatinine, and protein excretion. At the start of the study, fasting plasma glucose concentrations were measured. Intravenous insulin at 0.015 units·kg-1·h-1, titrated to maintain blood glucose between 80 and 150 mg/dl was given to participants with type 1 diabetes. In participants with type 2 diabetes, oral hypoglycaemic agents were witheld that morning, with those that required insulin receiving half of their usual morning dose of intermediate-acting insulin. After an 8 hour fast, individuals were studied in the supine position. An intravenous catheter for infusion and blood sampling was placed in each arm at 8 am. An automatic recording device measured blood pressure avery 15 minutes. To establish baseline renal haemodynamic measurements, participants were administered with a loading dose of 8 mg/kg of para-aminohippurate	Women with diabetes= 7.5±0.3 Women without diabetes= NA Plasma renin activity (ng Ang I·ml-1·h-1) Women with diabetes= 0.53±0.14 Women without diabetes= 0.52±0.14 Urine Na (mmol/24 hours) Women with diabetes= 270±28 Women without diabetes= 272±25 Urine protein (mg/24 hours) Women with diabetes= 94±44 Women without diabetes= 5±1 p<0.05 Microalbuminuria Women with diabetes= 6/9 (67%) Women without diabetes= 0/10 (0%) Glomerular filtration rate (ml·min-1·1.73 m-2) (median of readings at 10, 5, and 0 minutes before administration of oral captopril)	Yes 1.3 The prognostic factor of interest is adequately measured in study participants, sufficient to limit potential bias - Not applicable 1.4 The outcome of interest is adequately measured in study participants, sufficient to limit potential bias - Yes 1.5 Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest - Not applicable 1.6 The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of invalid results - Yes Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	These are characteristics for all of the women included in the study, including those who were not taking oral contraceptives. Women with diabetes taking oral contraceptives: Type I diabetes= 11/12 (92%) The type of diabetes in the remaining woman in this group was not specified in the study. The other characteristics of only the women taking oral contraceptives were not reported separately. Inclusion criteria Not reported Exclusion criteria Not reported		(PAH) and 50 mg/kg of inulin followed by constant infusions of PAH at 12mg/minute and inulin at 30 mg/minute for 90 minutes. This was followed by 25 mg captopril, taken orally. PAH clearance, inulin clearance, and plasma renin activity were measured at baseline. Serum PAH and inulin were measured by an autoanalyser. Plasma renin activity was assayed by radioimmunoassay. Urinary albumin concentration was measured by immunonephelometry. The baseline characteristics of the study participants were compared using non-parametric methods. Frequencies were comparing using the X2 test. An interaction between diabetes status and oral contraceptive use was checked using Friedman's test. Statistical analyses were performed with two-tailed significance levels of 0.05. *This study was performed	Women with diabetes= 129±4 Women without diabetes= 131±9 Renal plasma flow (ml·min-1·1.73 m-2) (median of readings at 10, 5, and 0 minutes before administration of oral captopril) Women with diabetes= 585±17 Women without diabetes= 623±30 Filtration fraction Women with diabetes= 0.22±0.01 Women without diabetes= 0.19±0.01 *This study was performed in two parts. The first was a comparative observational study comparing four groups of women: women with diabetes taking oral contraceptives, women without diabetes taking oral contraceptives, women with diabetes not taking oral contraceptives, and women without diabetes not taking oral contraceptives. The second part was an	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			in two parts. The first was a comparative observational study comparing four groups of women: women with diabetes taking oral contraceptives, women without diabetes taking oral contraceptives, women with diabetes not taking oral contraceptives, and women without diabetes not taking oral contraceptives. The second part was an intervention study on the use of captopril in women with diabetes. The intervention, methods and results for the first part of the study are reported here.	intervention study on the use of captopril in women with diabetes. The intervention, methods and results for the first part of the study are reported here.	
Full citation Tanis,B.C., van den Bosch,M.A., Kemmeren,J.M., Cats,V.M., Helmerhorst,F.M., Algra,A., van der,Graaf Y., Rosendaal,F.R., Oral contraceptives and the risk of myocardial infarction, New England Journal of MedicineN Engl J Med, 345, 1787-1793, 2001 Ref Id 216870	Sample size Whole study: n= 1173 (Myocardial infarction group= 248 Control group= 925) Subgroup of interest to NCC-WCH review: n= 446 (Women with diabetes= 7 Women without diabetes= 439) Characteristics	Interventions None	Details The study protocol was approved by the ethics committees of the participating hospitals. Oral informed consent was obtained from all participants. Participants in the myocardial infarction group were identified through a search of computerised hospital data bases. The International Classification of Diseases, 9th Revision,	Results Most of the results presented in the study paper compared factors in women who had a myocardial infarction and women who had not. Only the results for women who had used oral contraceptives are reported here. Women with diabetes: Myocardial infarction= 5/7 (71%) No myocardial infarction=	Limitations NICE guidelines manual. Appendix E: Methodology checklist: Case-control studies 1.1 The study addresses an appropriate and clearly focused question - Well covered 1.2 The cases and controls are taken from comparable populations -

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Country/ies where the study was carried out The Netherlands Study type Case-control study Aim of the study To investigate whether the use of low-dose combined oral contraceptives affects the risk of myocardial infarction Study dates January 1990 to October 1995 Source of funding Supported by a grant from the Netherlands Heart Foundation. One author had supervised research studies sponsored by multiple pharmaceutical companies that manufacture oral-contraceptive agents	Characteristics of all included women in the study (n=1173): Age (years): Myocardial infarction group= 42.7±6.5 (range 24 to 49) Control group= 38.1±8.3 (range 18 to 49) White ethnicity: Myocardial infarction group= 234/248 (94%) Control group= 864/925 (93%) Level of education: Primary school or less Myocardial infarction group= 130/247 (53%) Control group= 278/920 (30%) Secondary school Myocardial infarction group= 91/247 (37%) Control group= 390/920 (42%) Higher education or university Myocardial infarction group= 26/247 (11%) Control group= 252/920 (27%) (Level of education data missing for 1 woman with myocardial infarction and 5 controls) History of hypertension: Myocardial infarction group= 59/248 (24%) Control group= 56/921 (6%) (History of hypertension data missing for 4 controls) History of hypercholesterolaemia:		Clinical Modification codes for acute myocardial infarction were used. Participants in the control group were identified and recruited through random digit dialling. Telephone numbers were randomly generated by computer and then dialled until someone answered, or at least seven attempts had been made on different days and at different times of day. 15,725 telephone calls were made, 98% of the telephone numbers were answered. If a woman who was eligible to participate lived at the household contacted, she was asked to participate. Age differences between the myocardial infarction group and the control group were minimised by increasing the age limit of eligibility criteria during recruitment of controls. Controls were recruited from six geographic areas (based on where the women in the myocardial infarction group lived) and each control randomly received one of six questionnaires. The six forms of the questionnaire corresponded to the 6	2/7 (29%) Women without diabetes: Myocardial infarction= 94/439 (21%) No myocardial infarction= 345/439 (79%)	Adequately addressed 1.3 The same exclusion criteria are used for both cases and controls - Well covered 1.4 What was the participation rate for each group (cases and controls)? - 92% cases, 73% controls 1.5 Participants and non-participants are compared to establish their similarities or differences - Adequately addressed 1.6 Cases are clearly defined and differentiated from controls - Well covered 1.7 It is clearly established that controls are not cases - Well covered 1.8 Measures were taken to prevent knowledge of primary exposure from influencing case ascertainment - Not applicable 1.9 Exposure status is measured in a standard, valid, and reliable way - Well covered 1.10 The main potential confounders

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Myocardial infarction group= 28/248 (11%) Control group= 24/920 (3%) (History of hypercholesterolaemia data missing for 5 controls) History of diabetes: Myocardial infarction group= 15/248 (6%) Control group= 13/921 (1%) (History of diabetes data missing for 4 controls) Body Mass Index (BMI): Myocardial infarction group= 25.7±5.1 Control group= 23.5±3.9 (Body mass index data missing for 30 controls) Smoking status: Never smoked Myocardial infarction group= 21/248 (8%) Control group= 305/921 (33%) Former smoker Myocardial infarction group= 19/248 (8%) Control group= 222/921 (24%) Current smoker Myocardial infarction group= 208/248 (84%) Control group= 394/921 (43%) (Smoking status data missing for 4 controls) Family history of cardiovascular disease: Myocardial infarction group=		years in which women in the myocardial infarction group had been hospitalised for their first event. Therefore, the control group were a population sample stratified by age, geographical area, and calendar year. The questionnaires asked for information based on either the date of myocardial infarction (for participants in the myocardial infarction group), or the mid-year (for controls). Questions included body mass index, menopausal status, level of education, family history, history of hypertension, diabetes, hypercholesterolaemia, alcohol use, smoking, and the use of oral contraceptives. Women were classified as having hypertension, diabetes, or hypercholersterolaemia if they reported diagnosis by a clinician, or that they had been taking medication for the condition prior to the index date. A family history of cardiovascular disease was defined as the occurence of myocardial infarction, stroke, or		are identified and taken into account in the design and analysis - Adequately addressed 1.11 Have confidence intervals been provided? - Yes Other information Myocardial infarction was defined as the presence of symptoms, elevated cardiac-enzyme levels, and electrocardiographic changes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	156/239 (65%) Control group= 311/871 (36%) (Family history of cardiovascular disease data missing for 9 women with myocardial infarction and 54 controls) Premenopausal: Myocardial infarction group= 205/248 (83%) Control group= 767/925 (83%) Values are means±SD Characteristics were not presented separately for women with and without diabetes The study did not report the number of women with type 1 and type 2 diabetes Inclusion criteria Women in the myocardial infarction group: Women aged 18-49 years Women who were hospitalised for a first myocardial infarction between January 1990 and October 1995 Women in the control group: Women aged 18-49 years No history of coronory, cerebral, or peripheral arterial disease		peripheral arterial disease in at least one first-degree relative before the age of 60 years.		

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Exclusion criteria Women in the myocardial infarction group: Women who died during admission (n=19) Women who died between discharge and the start of the study (n= 9) Women who were 'unable to participate' (n=1) Women who could not be located (n= 21) Women who declined to participate (n= 23) Women who used oral contraceptive formulations other than those containing 30 µg of ethinyl estradiol Women in the control group: Women who used oral contraceptive formulations other than those containing 30 µg of ethinyl estradiol				
Full citation Diab,K.M., Zaki,M.M., Contraception in diabetic women: comparative metabolic study of Norplant, depot medroxyprogesterone acetate, low dose oral contraceptive pill and CuT380A, Journal of Obstetrics and Gynaecology Research, 26, 17-26, 2000	Sample size 40 women with diabetes* *85 women were recruited, but 40 women used either Norplant (n=20) or DMPA (n=20) - these are not relevant to the current review and so the results for these women are not reported here. 5 women changed their method of contraception during follow-up and were excluded from the	Interventions Oral contraceptive pill= 20 women Intrauterine contraceptive device= 20 women	Details Women were recruited from 'the Diabetic Institute.' All participants were counselled for different types of contraception. Women were included if they requested the use of an intrauterine contraceptive device (CuT380A IUD, FEI product, N Tonawanda	Results Systolic blood pressure (mmHg) (mean +/- standard error of the mean) Oral contraceptives group: At baseline= 113 +/- 0.99 3 months= 112 +/- 0.92 6 months= 112 +/- 0.52 9 months= 112 +/- 0.74 No significant difference between baseline and treatment values	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation to treatment groups was unrelated to potential confounding factors - Yes A2 Attempts were

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Ref Id 202828 Country/ies where the study was carried out Egypt Study type Prospective observational study Aim of the study To determine the long-term use of Norplant, depot medroxyprogesterone acetate (DMPA), and low dose oral contraceptives on glycaemic control, lipoprotein metabolism, and coagulation profile in women with diabetes. Study dates January 1996 to August 1997 Source of funding None reported	analysis - 1 woman in the IUD group had persistent vaginal bleeding, 1 woman in the Norplant group developed an infection where Norplant was implanted, 1 woman on oral contraceptives changes to IUD (no reason given), and 2 women in the DMPA group left the study after their first injection due to irregular vaginal bleeding. Characteristics Age (years) Range of total study sample= 20 to 40 Oral contraceptives group= 29.9 +/- 0.99 IUD group= 29.7 +/- 1.24 No significant differences between the groups were reported Age > 35 years Oral contraceptives group= 4/20 (20%) IUD group= 5/20 (25%) No significant differences between the groups were reported Women with type 1 diabetes Oral contraceptives group= 17/20 (85%) IUD group= 15/20 (75%) No significant differences between the groups were reported		USA), Levongestrel implant (6 silastic capsules each containing 36mg Levonorgestrel, Norplant, Leiras, Finland), depot medroxyprogresterone acetate (150mg, DMPA, Upjhon, USA), or the use of the low dose oral contraceptive pill (monophasic combination of 30 ug ethinyl estradiol and 75 ug gestodene, Gynera, Schering, Germany).* Informed consent was obtained. Women were able, after counselling, to chose which form of contraception they wished to use. Women were followed up at 3 months, 6 months, and 9 months. Women were asked about problems with the contraceptive method used at each follow up meeting. *40 women chose to use either Norplant or DMPA. These methods of contraception are not relevant to the current review and therefore the results for these women are not reported here.	IUD group: At baseline= 112 +/- 0.91 3 months= 110 +/- 0.50 6 months= 111 +/- 0.69 9 months= 111 +/- 0.50 No significant difference between baseline and treatment values Diastolic blood pressure (mmHg) (mean +/- standard error of the mean) Oral contraceptives group: At baseline= 73.5 +/- 1.31 3 months= 72.5 +/- 1.23 6 months= 72.0 +/- 1.17 9 months= 71.5 +/- 1.31 Value at 9 months is significantly different to baseline value IUD group: At baseline= 74.5 +/- 1.14 3 months= 71.0 +/- 1.00 6 months= 69.0 +/- 0.50 9 months= 67.5 +/- 0.99 Values at 3 months, 6 months, and 9 months are significantly different to baseline value Total cholesterol (mg/dl) (mean +/- standard error of the mean): Oral contraceptives group: At baseline= 209.2 +/- 6.57 3 months= 195.3 +/- 7.63 6 months= 205.5 +/- 6.67 9 months= 200.1 +/- 6.75	made within the design or analysis to balance the comparison groups for potential confounders - No A3 Groups were comparable at baseline, including all major confounding and prognostic factors - Yes B1 Comparison groups received the same care apart from the intervention(s) studied - Unclear B2 Participants receiving care were kept 'blind' to treatment allocation - N/A B3 Individuals administering care were kept 'blind' to treatment allocation - No C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? - None C2 b. Groups were comparable for

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Women with type 2 diabetes Oral contraceptives group= 3/20 (15%) IUD group= 5/20 (25%) No significant differences between the groups were reported Duration of diabetes (years) Oral contraceptives group= 6.15 +/- 1.10 IUD group= 5.35 +/- 1.03 No significant differences between the groups were reported HbA_{1c} (%) Oral contraceptives group= 6.95 +/- 0.17 IUD group= 7.10 +/- 0.16 No significant differences between the groups were reported BMI (kg/m2) Oral contraceptives group= 27.70 +/- 0.27 IUD group= 27.79 +/- 0.24 No significant differences between the groups were reported BMI > 27.5 Oral contraceptives group= 6/20 (30%) IUD group= 10/20 (50%) No significant differences			Value at 3 months is significantly different to baseline value IUD group: At baseline= 223.4 +/- 4.71 3 months= 211.1 +/- 5.68 6 months= 209.9 +/- 5.45 9 months= 218.1 +/- 4.96 Values at 3 months and 6 months are significantly different to baseline value Triglycerides (mg/dl) (mean +/- standard error of the mean) Oral contraceptives group: At baseline= 127.6 +/- 3.44 3 months= 137.9 +/- 2.74 6 months= 143.5 +/- 2.14 9 months= 148.9 +/- 2.29 Values at 3 months, 6 months, and 9 months are significantly different to baseline value IUD group: At baseline= 133.5 +/- 3.47 3 months= 138.8 +/- 5.23 6 months= 131.5 +/- 2.95 9 months= 136.2 +/- 3.79 No significant difference between baseline and treatment values High-density lipoprotein cholesterol (HDL-C, mg/dl) (mean +/- standard error of the mean)	treatment completion - Yes C3 a. For how many participants in each group were no outcome data available? - None C3 b. Groups were comparable with respect to the availability of outcome data - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - No D5 Investigators were kept 'blind' to other important confounding and prognostic factors - No Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	between the groups were reported All women were in stable glycemic control (HbA_{1c} less than 8%) All women were normotensive (systolic blood pressure less than 140, diastolic blood pressure less than 90) and had comparable systolic and diastolic blood pressure No women had evidence of diabetic complications as proliferative retinopathy or proteinuric nephropathy No women had current or past liver disease or thrombotic disorders All women were non-smokers All women had regular menstrual cycles None of the women received hormonal contraception for the 3 months prior to entry to the study None of the women had taken any medication known to interfere with haemostatic function, including salicylic acid, in the 4 weeks prior to entering the study *This includes 40 women who were using either Norplant or DMPA that are not relevant to the current review and therefore are not reported here			Oral contraceptives group: At baseline= 42.5 +/- 1.86 3 months= 54.8 +/- 3.79 6 months= 55.2 +/- 3.29 9 months= 56.3 +/- 3.35 Values at 3 months, 6 months, and 9 months are significantly different to baseline value IUD group: At baseline= 42.0 +/- 2.06 3 months= 44.2 +/- 2.79 6 months= 46.6 +/- 2.80 9 months= 41.2 +/- 1.40 No significant difference between baseline and treatment values Low density lipoprotein cholesterol (LDL-C, mg/dl) (mean +/- standard error of the mean) Oral contraceptives group: At baseline= 138.3 +/- 6.4 3 months= 129.1 +/- 6.97 6 months= 116.9 +/- 8.41 9 months= 107.3 +/- 5.85 Values at 3 months, 6 months, and 9 months are significantly different to baseline value IUD group: At baseline= 135.4 +/- 4.12 3 months= 129.1 +/- 5.70 6 months= 125.8 +/- 6.60 9 months= 132.9 +/- 3.40 No significant difference between baseline and treatment values	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Inclusion criteria None reported Exclusion criteria None reported			No significant change in the insulin or oral treatment dose among women in the different groups (actual data not reported) No differences seen in percentage change of mean weight and systolic blood presure (actual data not reported) Side effects varied with the chosen method of contraception, but none required a change of contraceptive method. 2 women in the IUD group had lower abdominal pain and vaginal discharge, one of whom also had menorrhagia. 2 women in the oral contraceptive group developed menstrual problems.	
Full citation Garg,S.K., Chase,H.P., Marshall,G., Hoops,S.L., Holmes,D.L., Jackson,W.E., Oral contraceptives and renal and retinal complications in young women with insulin-dependent diabetes mellitus, JAMA, 271, 1099-1102, 1994 Ref Id 203336	Sample size 86 women with diabetes Characteristics Age on first visit (years) (mean +/- standard error) Oral contraceptives group= 12.74 +/- 0.54 No oral contraceptives group= 13.72 +/- 0.68 P value not reported	Interventions None	Details All women signed a consent form approved by the University of Colorado Health Sciences Center Human Subjects Committee A power analysis was reported. The exact results were not reported, but it was reported that the sample size met the size required to detect differences of 0.6% for	Results HbA_{1c} (%) (mean +/- standard error of all years) Oral contraceptives group= 11.64 +/- 0.24 No oral contraceptives group= 11.86 +/- 0.24 P value not reported Cholesterol (mmol/L) (mean +/- standard error of all years) Oral contraceptives	Limitations NICE guidelines manual. Appendix E: Methodology checklist: Case-control studies 1.1 The study addresses an appropriate and clearly focused question - well covered 1.2 The cases and

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Country/ies where the study was carried out USA Study type Case-control study Aim of the study To determine whether oral contraceptives are a possible risk factor for early diabetic renal and/or retinal complications Study dates Not reported Source of funding None reported	Age on last visit (years) (mean +/- standard error) Oral contraceptives group= 22.69 +/- 0.46 No oral contraceptives group= 22.21 +/- 0.43 P value not reported Duration of diabetes on first visit (years) (mean +/- standard error) Oral contraceptives group= 3.90 +/- 0.63 No oral contraceptives group= 5.46 +/- 0.86 P value not reported Duration of diabetes on last visit (years) (mean +/- standard error) Oral contraceptives group= 13.84 +/- 0.77 No oral contraceptives group= 13.95 +/- 0.79 P value not reported Length of use of oral contraceptives in oral contraceptives group (years) (mean +/- standard deviation) Both groups= 3.4 +/- 2.9 (range 1.0 to 7.0 years) Inclusion criteria ≥ 14 years old Diabetes for ≥ 5 years Followed up in eye-kidney		HbA_{1c}, 0.31 mmol/L for cholesterol, change of 0.25 in eye grade, and 12.0 µg/min in albumin excretion rate. Out of 295 women who were included in a different study, this study used 43 women that met the inclusion criteria (oral contraceptives group). These 43 women were computer matched to 43 additional women by race, age, and duration of diabetes to serve as the comparison group. Women in the comparison group did not use oral contraceptives (no oral contraceptives group). Women in the oral contraceptives group were using various oral contraceptives and several reported changing their brands. All were using low-dose preparations containing 0.05mg or less of ethinyl estradiol (or mestranol) and a progestin. All women had direct ophthalmoscopy with pupils dilated by at least two examiners (an ophthalmologist and a diabetologist) followed by seven standard-field colour retinal photographs, intravenous fluorescein	group= 4.75 +/- 0.14 No oral contraceptives group= 4.64 +/- 0.11 P value not reported Diastolic blood pressure* Normal: Oral contraceptives group= 20/43 (47%) No oral contraceptives group= 20/43 (47%) No significant difference between groups (p=0.99) Borderline: Oral contraceptives group= 23/43 (53%) No oral contraceptives group= 23/43 (53%) No significant difference between groups (p=0.99) Systolic blood pressure* Normal: Oral contraceptives group= 31/43 (72%) No oral contraceptives group= 27/43 (63%) No significant difference between groups (p=0.36) Borderline: Oral contraceptives group= 12/43 (28%) No oral contraceptives group= 16/43 (37%) No significant difference between groups (p=0.36) Overnight albumin excretion rates on first visit	controls are taken from comparable populations - adequately covered 1.3 The same exclusion criteria are used for both cases and controls - well covered 1.4 What was the participation rate for each group (cases and controls)? - 100% 1.5 Participants and non-participants are compared to establish their similarities or differences - not applicable 1.6 Cases are clearly defined and differentiated from controls - well covered 1.7 It is clearly established that controls are not cases - well covered 1.8 Measures were taken to prevent knowledge of primary exposure from influencing case ascertainment - well covered 1.9 Exposure status is measured in a standard, valid, and reliable way - well covered 1.10 The main

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	clinic at least once a year Brought in a minimum of two overnight urine samples for albumin determinations Use of oral contraceptives for ≥ 1 year (for oral contraceptives group) Exclusion criteria Women who had ever been pregnant		photography (if necessary), and slit-lamp examinations. Retinal findings were graded using a modified Airlie House classification of diabetic retinopathy. The final eye grades for each eye were assigned in a maked fashion by one of the two retinal specialists based on the data of seven standard-field photographs. The category assigned to the worse eye was used for stastical analysis. Eye classifications were either normal (grade 1), background diabetic retinopathy (grades 2 to 4), preproliferative diabetic retinopathy (grade 5), or proliferative diabetic retinopathy (grade 6). A borderline elevated diastolic blood pressure or systolic blood pressure was reported if levels above the 90th percentile for age were found on at least two separate visits. Percentiles were taken from the Bogalusa Heart Study. All participants were asked to avoid caffeine, alcohol, and heavy exercise on the evenings of overnight urine specimen collections and	not reported Overnight albumin excretion rates on last visit (µg/min) < 7.6: Oral contraceptives group= 25/43 (58%) No oral contraceptives group= 28/43 (65%) No significant difference between groups (p=0.18) 7.6 to 20: Oral contraceptives group= 8/43 (19%) No oral contraceptives group= 9/43 (21%) No significant difference between groups (p=0.18) 20 to 200: Oral contraceptives group= 10/43 (23%) No oral contraceptives group= 4/43 (9%) No significant difference between groups (p=0.18) > 200: Oral contraceptives group= 0/43 (0%) No oral contraceptives group= 2/43 (5%) No significant difference between groups (p=0.18) Eye grades on last visit 1: Oral contraceptives group= 10/40 (25%) No oral contraceptives group= 6/39 (15%)	potential confounders are identified and taken into account in the design and analysis - adequately addressed 1.11 Have confidence intervals been provided? - not applicable Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			not to do a collection during menses or if a urinary tract infection was possibly present (all urine samples were analysed for leukocytes for a possible urinary tract infection - if they were present, the sample was discarded). The mean of two overnight urine samples was used for each eye-kidney visit.	No significant difference between groups (p=0.22) 2: Oral contraceptives group= 20/40 (50%) No oral contraceptives group= 16/39 (41%) No significant difference between groups (p=0.22) 3: Oral contraceptives group= 5/40 (13%) No oral contraceptives group= 12/39 (31%) No significant difference between groups (p=0.22) 4: Oral contraceptives group= 4/40 (10%) No oral contraceptives group= 7/39 (18%) No significant difference between groups (p=0.22) 5 to 6: Oral contraceptives group= 4/40 (10%) No oral contraceptives group= 2/39 (5%) No significant difference between groups (p=0.22) No change in eye grade Oral contraceptives group= 23/40 (58%) No oral contraceptives group= 23/39 (59%) No significant difference between groups (p=0.67) Worsening by 1 eye grade	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				Oral contraceptives group= 9/40 (23%) No oral contraceptives group= 8/39 (21%) No significant difference between groups (p=0.67) Worsening by > 1 eye grade Oral contraceptives group= 8/40 (20%) No oral contraceptives group= 6/39 (15%) No significant difference between groups (p=0.67) Improvement by 1 eye grade Oral contraceptives group= 0/40 (0%) No oral contraceptives group= 2/39 (5%) No significant difference between groups (p=0.67) Data unavailable for eye grade Oral contraceptives group= 3/43 (7%) No oral contraceptives group= 4/43 (9%) No significant difference between groups (p=0.67) No change in renal/microalbuminuria status Oral contraceptives group= 36/41 (88%)** No oral contraceptives	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				group= 35/40 (88%) Worsening of renal/microalbuminuria status (from 20.0 [normal] to 200.0 µg/min [microalbuminuria]) Oral contraceptives group= 5/41 (12%) No oral contraceptives group= 3/40 (8%) Improvement in renal/microalbuminuria status (not quantified) Oral contraceptives group= 0/41 (0%) No oral contraceptives group= 2/40 (5%) Data unavailable for change in renal/microalbuminuria status Oral contraceptives group= 2/43 (5%) No oral contraceptives group= 3/43 (7%)*** *It is not clear which visit this data was recorded from or whether it is a mean of all visits **One woman had macroalbuminuria and so her condition could not worsen ***Data was available for 40 women. The paper states that data was	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				unavailable for 2 women. The reviewer has assumed that data is also unavailable for the 1 woman who is unaccounted for.	
Full citation Grigoryan,O., Grodnitskaya,E., Andreeva,E., Shestakova,M., Melnichenko,G., Dedov,I., Contraception in perimenopausal women with diabetes mellitus, Gynecological Endocrinology, 22, 198-206, 2006 Ref Id 202830 Country/ies where the study was carried out Russia Study type Randomised controlled trial Aim of the study To assess how combined oral contraceptives and intrauterine devices affect carbohydrate and lipid metabolism and hemostasis in women	Sample size 153 women Characteristics Mean age 44.3 +/- 5.2 years Average age of onset: Type 1 diabetes= 24.6 +/- 4.9 years Type 2 diabetes= 38.1 +/- 2.8 Average duration: Type 1 diabetes= 14.3 +/- 3.8 years Type 2 diabetes= 5.3 +/- 4.7 years Non-proliferative retinopathy: Type 1 diabetes= 15 (26%) Type 2 diabetes= 39 (71%) Pre-proliferative retinopathy: Type 1 diabetes= 43 (74%) Type 2 diabetes= 16 (29%) Inclusion criteria Diabetes mellitus No evidence of proliferative	Interventions Combined low estrogen contraceptives= 28 women Combined standard dose contraceptives= 20 women Combined low progestogen contraceptives= 21 women Intrauterine contraceptive device= 22 women No contraceptives= 40 women An additional group of 22 women (11 type 1 diabetes, 11 type 2 diabetes) were given a progestogen intrauterine contraceptive device (Mirena LNG-IUS - Schering, Germany) but the results of this group are not reported here as they are not relevant to the NCC-WCH review.	Details The study protocol and informed consent documents were approved by the local ethics committee. All women gave signed informed consent before participating in the study. Before the study started, women were randomised using a computer-generated scheme to one of five treatment groups or the control group: One group consisted of 28 women (14 type 1 diabetes, 14 type 2 diabetes) who were given a pill of 20µg ethinylestradiol and 150µg desogestrel (Novinet - Gedeon Richter, Hungary) (combined low oestrogen oral contraceptives group). A second group consisted of 20 women (10 type 1 diabetes, 10 type 2 diabetes) who were given a pill of 30µg ethinylestradiol and 150µg	Results HbA_{1c} (%) Combined low estrogen contraceptives group Type 1: At baseline= 7.5 +/- 0.3 3 months= 7.6 +/- 0.5 6 months= 7.4 +/- 0.4 9 months= 7.6 +/- 0.3 12 months= 7.5 +/- 0.6 No significant differences reported Type 2: At baseline= 7.6 +/- 0.5 3 months= 7.5 +/- 0.6 6 months= 7.7 +/- 0.3 9 months= 7.4 +/- 0.5 12 months= 7.5 +/- 0.7 No significant differences reported Combined standard dose contraceptives group Type 1: At baseline= 7.5 +/- 0.3 3 months= 7.6 +/- 0.2 6 months= 7.4 +/- 0.4 9 months=7.6 +/- 0.6 12 months= 7.5 +/- 0.4 No significant differences reported	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials A1 An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups) - unclear A2 There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation) - unclear A3 The groups were comparable at baseline, including all major confounding

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
with diabetes Study dates November 2002 to July 2003 Source of funding None reported	retinopathy, nephropathy or macrovascular complications Exclusion criteria 'Type 1 and type 2 diabetes mellitus women in the state of decompensation of the primary disease' Ketoacidosis A history of myocardial infarction and/or thromboembolism during the year prior to the start of the study Elevated blood creatinine and urea Nodular form of fibrous-cystic mastopathy The presence of any oncological diseases at the time of study Lack of self-control skills Smokers		desogestrel (Marvelon - Organon, The Netherlands) (combined standard dose oral contraceptives group). A third group consisted of 21 women (12 type 1 diabetes, 9 type 2 diabetes) who were given a pill of 30µg ethinylestradiuol and 75µg gestodene (combined low progestogen oral contraceptives group). A fourth group consisted of 22 women (11 type 1 diabetes, 11 type 2 diabetes) who were given a copper-containing intrauterine contraceptive device (intrauterine device group). A fifth group consisted of 40 aged-matched controls who did not use any methods of contraception (no contraceptives group). A sixth group of 22 women (11 type 1 diabetes, 11 type 2 diabetes) were given a progestogen intrauterine contraceptive device (Mirena LNG-IUS - Schering, Germany) but the results of this group are not reported here as they are not relevant to the NCC-WCH review. All women enrolled in the study completed the	Type 2: At baseline= 7.7 +/- 0.4 3 months= 7.8 +/- 0.5 6 months= 7.6 +/- 0.7 9 months= 7.5 +/- 0.4 12 months= 7.6 +/- 0.3 No significant differences reported Combined low progestogen contraceptives group Type 1: At baseline= 7.5 +/- 0.3 3 months= 7.6 +/- 0.2 6 months= 7.4 +/- 0.4 9 months= 7.6 +/- 0.6 12 months= 7.5 +/- 0.4 No significant differences reported Type 2: At baseline= 7.3 +/- 0.4 3 months= 7.4 +/- 0.6 6 months= 7.5 +/- 0.5 9 months= 7.6 +/- 0.3 12 months= 7.4 +/- 0.7 No significant differences reported Intrauterine device group Type 1: At baseline= 7.8 +/- 0.3 3 months= 7.7 +/- 0.8 6 months= 7.9 +/- 0.2 9 months= 7.5 +/- 0.6 12 months= 7.8 +/- 0.7 No significant differences reported Type 2: At baseline= 7.5 +/- 0.7 3 months= 7.7 +/- 0.4 6 months= 7.5 +/- 0.7 9 months= 7.6 +/- 0.4	and prognostic factors - unclear B1 The comparison groups received the same care apart from the intervention(s) studied - unclear B2 Participants receiving care were kept 'blind' to treatment allocation - unclear B3 Individuals administering care were kept 'blind' to treatment allocation - unclear C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - yes C2 a. How many participants did not complete treatment in each group? - none C2 b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment) - yes C3 a. For how many participants in each

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			study. Women who eliminated their intrauterine device were not excluded from the statistical analysis. All women underwent a general clinical examination and gynecological examinations. Clinical and laboratory exminations were carried out at baseline and after 3, 6, 9 and 12 months of receiving contraception Findings were reported as significant if p < 0.05	12 months= 7.4 +/- 0.3 No significant differences reported No contraceptives group At baseline= 7.7 +/- 0.6 3 months= 7.5 +/- 0.3 6 months= 7.7 +/- 0.5 9 months= 7.6 +/- 0.7 12 months= 7.5 +/- 0.2 No significant differences reported Some data was reported for lipid levels, but not enough to allow a comparison between women using oral contraceptives and women not using oral contraceptives and so it is not reported here It is reported that the incidence of side effects in women using oral contraceptives and in women using intrauterine devices was not different from those seen in apparently healthy women In the oral contraceptives groups, 13 (19%) women (11 [31%] women with type 1, 2 [6%] women with type 2) had no side effects. Reported side effects were intermenstrual bloody discharge (7 [19%] women	group were no outcome data available? - none C3 b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available) - yes D1 The study had an appropriate length of follow-up - yes D2 The study used a precise definition of outcome - yes D3 A valid and reliable method was used to determine the outcome - yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - unclear D5 Investigators were kept 'blind' to other important confounding and prognostic factors - unclear Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				with type 1, 3 [9%] women with type 2), breast enlargement and tenderness (16 [44%] women with type 1, 10 [30%] women with type 2), gnawing pain in the lower limbs (5 [14%] women with type 1, 5 [15%] women with type 2), pain in the dextral hypochondrium (2 [6%] women with type 1, 4 [12%] women with type 2), and vaginal discharge (27 [75%] women with type 1, 15 [46%] women with type 2). The side effects data was not separated for the type of intrauterine device, and so the following data includes women in the progestogen intrauterine contraceptive device group who were excluded from the NCC-WCH review. In the intrauterine contraceptives groups, 5 (23%) women with type 1 and 4 (18%) women with type 2 diabetes had menstrual cycle disorders (including polymenorrhea, meno- and/or metrorrhagia), and 3 (14%) women with type 1 and 3 (14%) women with type 2 diabetes had pain syndrome. There was no	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				significant difference in the incidence and type of adverse effects in women with type 1 diabetes and women with type 2 diabetes. The intrauterine device was removed in 4 (18%) women with type 1 diabetes and 2 (9%) women with type 2 diabetes after 6 months due to persistent, frequent, intermenstrual bloody discharge. 2 (9%) women with type 1 diabetes had incomplete expulsion of the intrauterine device. There were no reported cases of inflammatory diseases of the small pelvis organs*. *This is the terminology used in the study paper	
Full citation Klein,B.E., Moss,S.E., Klein,R., Oral contraceptives in women with diabetes, Diabetes Care, 13, 895-898, 1990 Ref Id	Sample size 384 women Characteristics Age (years) 14 to 24= 110/384 (29%) 25 to 34= 138/384 (36%)	Interventions None	Details Informed consent was obtained from all women Physical and ocular examinations were performed on all women, including measuring blood pressure, dilating the	Results The authors of the paper state that the 'mild to minimal' and 'moderate to severe' categories of diabetic retinopathy are nonproliferative.	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation to treatment

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
203335 Country/ies where the study was carried out USA Study type Prospective observational study Aim of the study To investigate the relationship between oral contraceptive use and severity of diabetic retinopathy Study dates 1984 to 1986 Source of funding Supported by grants from the Retina Research Foundation (B.E.K.K.) and the National Eye Institute (EY-03083; R.K.)	35+ = 136/384 (35%) Use of birth control pills Never= 214/384 (56%) Ever= 170/384 (44%) Duration of use of birth control pills ≤ 1 year= 62/384 (16%) 2 to 4 years= 59/384 (15%) ≥ 5 years= 49/384 (13%) Using birth control pills at the time of this study Yes= 33/384 (9%) No= 351/384 (91%) Inclusion criteria At least 14 years old Women who take insulin Birth control pill history available Exclusion criteria None reported		pupils, taking stereoscopic fundus photographs of seven standard fields of each eye, determining blood glucose, determining glycosylated hemoglobin. A structured interview was used to determine whether the women had ever taken birth control pills, and if they had, the names and duration of use of the medications. Grading of retinopathy took place at the University of Wisconsin Fundus Photograph Reading Centre using the Early Treatment Diabetic Retinopathy Study adaptation of the moedified Airlie House classification of diabetic retinopathy, which was further adapted in-house. The level of retinopathy was determined by the most severely involved eye. For each eye, the maximum grade in any of the seven standard photographic fields was determined for each of the lesions and used in defining the retinopathy levels. The scale ranged from no retinopathy to the most severe retinopathy (including severe virteous hemorrhage, phthtisis	Never used birth control pills* No diabetic retinopathy= 31/214 (14%) Mild to minimal diabetic retinopathy= 88/214 (41%) Moderate to severe retinopathy= 43/214 (20%) Proliferative retinopathy= 52/214 (24%) Ever used birth control pills No diabetic retinopathy= 14/170 (8%) Mild to minimal diabetic retinopathy= 77/170 (45%) Moderate to severe retinopathy= 37/170 (22%) Proliferative retinopathy= 42/170 (25%) ≤ 1 year of use of birth control pills (excluding never used) No diabetic retinopathy= 6/62 (10%) Mild to minimal diabetic retinopathy= 25/62 (40%) Moderate to severe retinopathy= 16/62 (26%) Proliferative retinopathy= 15/62 (24%) 2 to 4 years of use of birth control pills No diabetic retinopathy= 4/59 (7%) Mild to minimal diabetic retinopathy= 33/59 (56%)	groups was unrelated to potential confounding factors - yes A2 Attempts were made within the design or analysis to balance the comparison groups for potential confounders - no A3 Groups were comparable at baseline, including all major confounding and prognostic factors - unclear B1 Comparison groups received the same care apart from the intervention(s) studied - unclear B2 Participants receiving care were kept 'blind' to treatment allocation - no B3 Individuals administering care were kept 'blind' to treatment allocation - unclear C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - yes C2 a. How many

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			bulbi, or enucleation). Hypertension is defined as ≥ 160 mmHg systolic and/or ≥95 mmHg diastolic for women aged 25 years or older, and ≥ 140 mmHg systolic and/or ≥ 90 mmHg diastolic for those aged under 25 years or on hypertension medication.	Moderate to severe retinopathy= 10/59 (17%) Proliferative retinopathy= 12/59 (20%) ≥ 5 years of use of birth control pills No diabetic retinopathy= 4/49 (8%) Mild to minimal diabetic retinopathy= 19/49 (39%) Moderate to severe retinopathy= 11/49 (22%) Proliferative retinopathy= 15/49 (31%) Currently using birth control pills No diabetic retinopathy= 39/351 (11%) Mild to minimal diabetic retinopathy= 147/351 (42%) Moderate to severe retinopathy= 74/351 (21%) Proliferative retinopathy= 91/351 (26%) Not currently using birth control pills No diabetic retinopathy= 6/33 (18%) Mild to minimal diabetic retinopathy= 18/33 (55%) Moderate to severe retinopathy= 6/33 (18%) Proliferative retinopathy= 3/33 (9%) No evidence of effect of	participants did not complete treatment in each group? - none C2 b. Groups were comparable for treatment completion - yes C3 a. For how many participants in each group were no outcome data available? - none C3 b. Groups were comparable with respect to the availability of outcome data - yes D1 The study had an appropriate length of follow-up - yes D2 The study used a precise definition of outcome - yes D3 A valid and reliable method was used to determine the outcome - yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - unclear D5 Investigators were kept 'blind' to other important confounding and prognostic factors - unclear Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				ever using birth control pills on severity of retinopathy when controlling individually for: current age, duration of diabetes, systolic blood pressure, diastolic blood pressure, glycosylated haemoglobin, proteinuria, or body mass index The following factors were significantly associated with the severity of retinopathy (ordinal logistic model): duration of diabetes, diastolic blood pressure, proteinuria, and glycosylated haemoglobin. Current use of birth control pills, prior use of birth control pills, and the years of use of birth control pills did not add significantly to the factors that were found to be significantly associated with the severity of retinopathy (no actual data reported). The following factors were not significantly associated with the severity of retinopathy: age, systolic blood pressure, and body mass**. Current use of birth control pills, prior use of birth	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				control pills, and the years of use of birth control pills did not add significantly to factors significantly associated with hypertension (no actual data reported) Current use of birth control pills, prior use of birth control pills, and the years of use of birth control pills was not significantly associated with glycosylated haemoglobin (no actual data reported) *Percentages do not add up to 100 due to rounding **It is not clear whether this refers to weight or body mass index	
Full citation Petersen,K.R., Skouby,S.O., Vedel,P., Haaber,A.B., Hormonal contraception in women with IDDM. Influence on glycometabolic control and lipoprotein metabolism, Diabetes Care, 18, 800-806, 1995 Ref Id 203099 Country/ies where the study was carried out Denmark	Sample size 42* women *1 woman in the oral contraceptives group and 2 women in the no oral contraceptives group withdrew from the study before baseline values were obtained. These women were not replaced. 5 women in the oral contraceptives group did not complete the study period - 2 left the study for personal reasons, 1 left due to	Interventions Women with diabetes using oral contraceptives (n= 22) Women with diabetes not using oral contraceptives (n= 20)	Details Informed consent was obtained from all participants. The study was approved by the Medical Ethics Committee of Copenhagen and the Danish National Board of Health. Women were recruited from attending an outpatient clinic for contraceptive counselling. Women were recruited	Results Arterial blood pressure (mmHg) (mean) Oral contraceptives group: Baseline= 90 (range 80 to 103) 12 months= 92 (range 79 to 109) P value not reported No oral contraceptives group: Baseline= 97 (range 75 to 113) 12 months= 94 (range 81 to 111)	Limitations NICE guidelines manual. Appendix E: Methodology checklist: Case-control studies 1.1 The study addresses an appropriate and clearly focused question - well covered 1.2 The cases and controls are taken from comparable

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Study type Case-control study Aim of the study To investigate the effect of long-term intake of oral contraceptives on glycemia control and lipoprotein metabolism Study dates Not reported Source of funding 'This study was supported by a grant from the Ove Villiam Buhl Olesen and Edith Buhl Olsesn Foundation and by the pharmaceutical company Schering, Denmark'	increased frequency and severity of hypoglycaemic attacks, and 2 left after 6 months due to abdominal discomfort and nausea. 1 woman from the no oral contraceptives group conceived after 4 months. Characteristics Age (years) (median) Oral contraceptives group= 26.5 (range 19 to 32) No oral contraceptives group= 28.5 (range 21 to 33) P reported as significant Duration of diabetes (years) (median) Oral contraceptives group= 9.5 (range 3 to 22) No oral contraceptives group= 11.5 (range 2 to 25) P not reported Smokers (less than 10 cigarettes a day) Oral contraceptives group= 11/22 (50%) No oral contraceptives group= 9/20 (45%) P not reported Arterial blood pressure (mmHg) (mean) Oral contraceptives group= 90 (range 80 to 103) No oral contraceptives group= 97 (range 75 to 113)		into the study if they wanted to use oral contraception. If oral contraceptives had been used previously, a 3 month washout period was used. The women received a monophasic combination of 30 µg ethinyl estradiol and 75 µg gestodene for 21 days, and then had 7 days free of medication, for 12 cycles. A control was selected for each participant in the oral contraception group - a woman of similar age, diabetic status, smoking habits, body mass index (BMI), marital and socioeconomic status using nonhormonal contraception The study authors aimed to recruit at least 17 women in each group to allow the smallest difference in baseline characteristics not to be overlooked	P value not reported HbA_{1c} (%) (median) Oral contraceptives group: Baseline= 8.2 (range 5.8 to 11.2) 12 months= 8.4 (range 6.0 to 10.8) P value not reported No oral contraceptives group: Baseline= 8.5 (range 6.4 to 11.7) 12 months= 8.2 (range 7.3 to 11.0) P value not reported Microalbuminuria (number of women) Oral contraceptives group: Baseline= 2/22 (9%) 12 months= 2/22 (9%) P value not reported No oral contraceptives group: Baseline= 3/20 (15%) 12 months= 2/20 (10%) P value not reported Free fatty acids (mmol/l) (median) Oral contraceptives group: Baseline= 0.88 (range 0.16 to 2.40) 12 months= 0.86 (range 0.22 to 1.42) P value not reported No oral contraceptives group: Baseline= 0.89 (range	populations - well covered 1.3 The same exclusion criteria are used for both cases and controls - well covered 1.4 What was the participation rate for each group (cases and controls)? - 100% 1.5 Participants and non-participants are compared to establish their similarities or differences - not applicable 1.6 Cases are clearly defined and differentiated from controls - well covered 1.7 It is clearly established that controls are not cases - well covered 1.8 Measures were taken to prevent knowledge of primary exposure from influencing case ascertainment - well covered 1.9 Exposure status is measured in a standard, valid, and reliable way - well covered 1.10 The main potential confounders are identified and

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	P not reported BMI (kg/m2) (median) Oral contraceptives group= 22.5 (range 19.1 to 25.4) No oral contraceptives group= 22.7 (range 17.9 to 31.6) P not reported Fasting plasma glucose (mmol/l) (median) Oral contraceptives group= 9.9 (range 1.8 to 19.7) No oral contraceptives group= 10.5 (range 5.2 to 22.6) P not reported HbA_{1c} (%) (median) Oral contraceptives group= 8.2 (range 5.8 to 11.2) No oral contraceptives group= 8.5 (range 6.4 to 11.7) P not reported 24 hour blood glucose level (mmol/l) (median) Oral contraceptives group= 8.7 (range 4.2 to 16.9) No oral contraceptives group= 7.5 (range 5.4 to 13.3) P not reported Daily insulin requirement (IU) (median) Oral contraceptives group= 41 (range 22 to 70) No oral contraceptives group= 42 (range 16 to 58) P not reported			0.32 to 2.52) 12 months= 1.11 (range 0.53 to 1.69) P value not reported Total cholesterol (mmol/l) (median) Oral contraceptives group: Baseline= 4.93 (range 3.06 to 7.97) 1 month= 4.64 (range 3.19 to 6.32) 3 months= 4.64 (range 3.44 to 7.51) 6 months**= 4.74 (range 3.10 to 6.93) 12 months***= 4.53 (range 3.09 to 6.52) No significant difference between baseline and any treatment values No oral contraceptives group: Baseline= 5.40 (range 3.46 to 7.08) 1 month= 5.23 (range 4.07 to 8.42) 3 months= 5.14 (range 4.28 to 8.03) 6 months**= 5.27 (range 4.05 to 7.56) 12 months**= 5.06 (range 3.77 to 7.45) No significant difference between baseline and any treatment values Low-density lipoprotein cholesterol (mmol/l) (median) Oral contraceptives group:	taken into account in the design and analysis - not reported 1.11 Have confidence intervals been provided? - not applicable Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Microalbuminuria (number of women) Oral contraceptives group= 2/22 (9%) No oral contraceptives group= 3/20 (15%) P not reported Free fatty acids (mmol/l) (median) Oral contraceptives group= 0.88 (range 0.16 to 2.40) No oral contraceptives group= 0.89 (range 0.32 to 2.52) P not reported Total cholesterol (mmol/l) (median) Oral contraceptives group= 4.93 (range 3.06 to 7.97) No oral contraceptives group= 5.40 (range 3.46 to 7.08) P not reported LDL cholesterol (mmol/l) (median) Oral contraceptives group= 3.16 (range 1.41 to 6.37) No oral contraceptives group= 3.27 (range 1.47 to 5.11) P not reported HDL cholesterol (mmol/l) (median) Oral contraceptives group= 1.36 (range 0.95 to 2.12) No oral contraceptives group= 1.64 (range 1.08 to 2.33)			Baseline= 3.16 (range 1.41 to 6.37) 1 month= 2.56 (range 0.98 to 4.52) 3 months= 2.55 (range 1.11 to 4.60) 6 months**= 2.55 (range 0.52 to 4.83) 12 months***= 2.46 (range 0.92 to 4.44) Values at 6 months and 12 months are significantly different to baseline value No oral contraceptives group: Baseline= 3.27 (range 1.47 to 5.11) 1 month= 3.24 (range 1.71 to 6.46) 3 months= 3.23 (range 2.01 to 5.21) 6 months**= 3.14 (range 1.79 to 5.71) 12 months**= 2.86 (range 1.81 to 4.71) Value at 12 months is significantly different to baseline value High-density lipoprotein cholesterol (mmol/l) (median) Oral contraceptives group: Baseline= 1.36 (range 0.95 to 2.12) 1 month= 1.43 (range 1.11 to 2.07) 3 months= 1.47 (range 0.88 to 1.98) 6 months**= 1.47 (range 1.06 to 2.13)	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	P not reported HDL2 cholesterol (mmol/l) (median) Oral contraceptives group= 0.64 (range 0.14 to 1.22) No oral contraceptives group= 0.86 (range 0.17 to 1.23) P not reported HDL3 cholesterol (mmol/l) (median) Oral contraceptives group= 0.75 (range 0.52 to 1.03) No oral contraceptives group= 0.83 (range 0.67 to 1.13) P not reported HDL cholesterol/total cholesterol (median) Oral contraceptives group= 0.31 (range 0.13 to 0.50) No oral contraceptives group= 0.31 (range 0.17 to 0.49) P not reported VLDL cholesterol (mmol/l) (median) Oral contraceptives group= 0.41 (range 0.18 to 2.76) No oral contraceptives group= 0.44 (range 0.26 to 0.84) P not reported Triglycerides (mmol/l) (median) Oral contraceptives group= 0.88 (range 0.39 to 5.98) No oral contraceptives group= 0.96 (range 0.56 to 1.83)			12 months***= 1.52 (range 1.14 to 2.21) No significant difference between baseline and any treatment values No oral contraceptives group: Baseline= 1.64 (range 1.08 to 2.33) 1 month= 1.70 (range 0.88 to 2.20) 3 months= 1.76 (range 0.89 to 2.20) 6 months**= 1.67 (range 0.99 to 2.13) 12 months**= 1.85 (range 0.88 to 2.75) No significant difference between baseline and any treatment values High-density lipoprotein2 cholesterol (mmol/l) (median) Oral contraceptives group: Baseline= 0.64 (range 0.14 to 1.22) 1 month= 0.67 (range 0.25 to 1.09) 3 months= 0.59 (range 0.11 to 1.17) 6 months**= 0.67 (range 0.20 to 1.23) 12 months***= 0.50 (range 0.20 to 1.18) No significant difference between baseline and any treatment values No oral contraceptives group: Baseline= 0.86 (range	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	P not reported No women had vascular or renal symptoms or had previously suffered from liver disease or thromboembolic disorders All women were at least six months postpartum or 3 months postabortem. None were lactating. Median values for age and duration of diabetes were similar in smokers and non-smokers. There was no significant difference in baseline values for smokers and non-smokers in the oral contraceptives group. It was not reported whether there was a significant difference or not in the no oral contraceptives group. Inclusion criteria Type 1 diabetes for at least 2 years and stable glycaemic control Exclusion criteria Smokers of 10 or more cigarettes a day			0.17 to 1.23) 1 month= 0.84 (range 0.07 to 1.37) 3 months= 0.83 (range 0.08 to 1.57) 6 months**= 0.92 (range 0.17 to 1.39) 12 months**= 0.88 (range 0.11 to 1.95) No significant difference between baseline and any treatment values High-density lipoprotein3 cholesterol (mmol/l) (median) Oral contraceptives group: Baseline= 0.75 (range 0.52 to 1.03) 1 month= 0.80 (range 0.59 to 1.10) 3 months= 0.86 (range 0.63 to 1.15) 6 months**= 0.88 (range 0.60 to 1.12) 12 months***= 1.00 (range 0.84 to 1.19) Values at 1 month, 3 months, 6 months and 12 months are significantly different to baseline value No oral contraceptives group: Baseline= 0.83 (range 0.67 to 1.13) 1 month= 0.84 (range 0.59 to 1.13) 3 months= 0.83 (range 0.63 to 1.11) 6 months**= 0.82 (range 0.69 to 1.15)	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				12 months**= 0.94 (range 0.70 to 1.30) Value at 12 months is significantly different to baseline value High-density lipoprotein cholesterol/total cholesterol (median) Oral contraceptives group: Baseline= 0.31 (range 0.13 to 0.50) 1 month= 0.33 (range 0.18 to 0.58) 3 months= 0.33 (range 0.15 to 0.53) 6 months**= 0.33 (range 0.19 to 0.69) 12 months***= 0.34 (range 0.18 to 0.57) No significant difference between baseline and any treatment values No oral contraceptives group: Baseline= 0.31 (range 0.17 to 0.49) 1 month= 0.32 (range 0.16 to 0.49) 3 months= 0.33 (range 0.16 to 0.46) 6 months**= 0.31 (range 0.14 to 0.50) 12 months**= 0.35 (range 0.17 to 0.59) Value at 12 months is significantly different to baseline value Very low density lipoprotein cholesterol	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				(mmol/l) (median) Oral contraceptives group: Baseline= 0.41 (range 0.18 to 2.76) 1 month= 0.47 (range 0.26 to 1.12) 3 months= 0.56 (range 0.26 to 0.88) 6 months**= 0.53 (range 0.39 to 2.01) 12 months***= 0.51 (range 0.40 to 1.67) Values at 1 month, 3 months, 6 months, and 12 months are significantly different to baseline value No oral contraceptives group: Baseline= 0.44 (range 0.26 to 0.84) 1 month= 0.43 (range 0.19 to 0.83) 3 months= 0.40 (range 0.22 to 1.00) 6 months**= 0.42 (range 0.29 to 1.10) 12 months**= 0.43 (range 0.29 to 1.16) No significant difference between baseline and treatment values Triglycerides (mmol/l) (median) Oral contraceptives group: Baseline= 0.88 (range 0.39 to 5.98) 1 month= 1.03 (range 0.57 to 2.43) 3 months= 1.23 (range 0.57 to 1.92)	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				6 months**= 1.14 (range 0.84 to 4.37) 12 months***= 1.10 (range 0.86 to 3.61) Values at 1 month, 3 months, 6 months, and 12 months are significantly different to baseline value No oral contraceptives group: Baseline= 0.96 (range 0.56 to 1.83) 1 month= 0.92 (range 0.41 to 1.81) 3 months= 0.87 (range 0.47 to 2.18) 6 months**= 0.92 (range 0.64 to 2.39) 12 months**= 0.94 (range 0.64 to 2.51) No significant difference between baseline and treatment values **Includes data for 19 women *** Includes data for 17 women	
Full citation Skouby,S.O., Molsted-Pedersen,L., Kuhl,C., Bennet,P., Oral contraceptives in diabetic women: metabolic effects of four compounds with different estrogen/progestogen	Sample size 27 women Characteristics Age (years) (mean +/- standard error) 22 +/- 3 (range 17 to 30 years) Age at onset of diabetes	Interventions Monophasic combined (high dose) group = 10 women* Monophasic combined (low dose) group = 10 women* Progesterone only group = 9 women* Triphasic combined group	Details Women who wanted to use oral contraceptives were recruited into the study. The study was approved by the local ethics committee and all participants gave informed consent.	Results HbA_{1c} (%) (assumed to be reported as mean +/- standard deviation) Monophasic combined high dose group: Baseline= 8.6 +/- 0.7 2 months= 9.4 +/- 0.6 6 months= 8.8 +/- 0.4 No significant difference	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials A1 An appropriate method of randomisation was used to allocate

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
profiles, Fertility and Sterility, 46, 858-864, 1986 Ref Id 203334 Country/ies where the study was carried out Denmark Study type Prospective randomised trial Aim of the study To compare the influence on metabolic effects and diabetes control of four different types of oral contraceptives Study dates Not reported Source of funding Supported by The Danish Diabetes Association and a grant from the Ove Villiam Buhl Olesen and Edith Buhl Olesen Memorial Foundation Contraceptive compounds (Triquilar, Microplan, and Gestaplan) were provided by Schering, Denmark and DAK Laboratories,	(years) (mean +/- standard error) 14 +/- 1.6 (range 1 to 19 years) HbA_{1c} (%) (assumed to be reported as mean +/- standard deviation) Monophasic combined (high dose) group= 8.6 +/- 0.7 Monophasic combined (low dose) group= 9.5 +/- 0.7 Progesterone only group= 8.9 +/- 0.5 Triphasic combined (low dose) group= 9.1 +/- 0.5 No significant difference between the three groups (p value not reported) All women had comparable socio economic status None of the women had used hormonal contraceptives for at least 6 weeks before entering the study No significant differences in mean body weight between the groups Inclusion criteria Women with insulin-dependent diabetes "Weight within 20% of ideal" Age < 35 years No evidence of late diabetic complications (e.g. background retinopathy or nephropathy	= 9 women* *After the first six months, 8 of the 27 women had a washout period of 6 weeks and then changed to one or more of the other groups, so the total number of women in the groups is larger than the sample size	Women were assigned to one of four groups at random (method of randomisation not reported). One group received a monophasic combination of tablets containing 4mg of 17β-estradiol (E2), 2mg of estradiol, and 3mg of norethindrone (monophasic combined high dose group). A second group received a combination of 35µg ethinyl E2 (EE2) and 500µg of norethindrone (monophasic combined low dose group). A third group received 300µg of norethindrone (progesterone only group). A fourth group received a combination of 30µg of EE2 + 50µg of levonorgestrel for the first 6 days, 40µg of EE2 + 75µg of levonorgestrel for the next 5 days, and 30µg of EE2 + 125µg of levonorgestrel during the last 10 days for each treatment cycle (triphasic combined group). All treatment regimens were given in six month periods. The three combined groups took their assigned medication for three weeks, followed	between baseline and treatment values (p values not reported) Monophasic combined low dose group: Baseline= 9.5 +/- 0.7 2 months= 8.2 +/- 0.3 6 months= 9.1 +/- 0.7 No significant difference between baseline and treatment values (p values not reported) Progesterone only group: Baseline= 8.9 +/- 0.5 2 months= 7.4 +/- 0.9 6 months= 9.5 +/- 0.9 No significant difference between baseline and treatment values (p values not reported) Triphasic combined group: Baseline= 9.1 +/- 0.5 2 months= 9.0 +/- 0.5 6 months= 9.1 +/- 0.5 No significant difference between baseline and treatment values (p values not reported) Free fatty acids (mmol/l) (assumed to be reported as mean +/- standard deviation) Monophasic combined high dose group: Baseline= 986 +/- 151 2 months= 814 +/- 100 6 months= 1033 +/- 145 No significant difference between baseline and treatment values (p values	participants to treatment groups (which would have balanced any confounding factors equally across groups) - unclear A2 There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation) - unclear A3 The groups were comparable at baseline, including all major confounding and prognostic factors - yes B1 The comparison groups received the same care apart from the intervention(s) studied - unclear B2 Participants receiving care were kept 'blind' to treatment allocation - unclear B3 Individuals administering care were kept 'blind' to treatment allocation - unclear C1 All groups were followed up for an

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Copenhagen, Denmark.	[serum creatinine < 120 nmol/l and blood pressure < 140/90]) Exclusion criteria None reported		by a week with no medication. The progesterone only group took their medication daily during the whole treatment period. Measurements were taken before treatment started and again after 2 months and 6 months of treatment. After the first six months, 8 of the 27 women had a washout period of 6 weeks and then changed to one or more of the other groups.	not reported) Monophasic combined low dose group: Baseline= 854 +/- 99 2 months= 996 +/- 112 6 months= 756 +/- 118 No significant difference between baseline and treatment values (p values not reported) Progesterone only group: Baseline= 969 +/- 138 2 months= 1030 +/- 251 6 months= 783 +/- 123 No significant difference between baseline and treatment values (p values not reported) Triphasic combined group: Baseline= 594 +/- 61 2 months= 452 +/- 151 6 months= 761 +/- 105 No significant difference between baseline and treatment values (p values not reported) Triglycerides (mmol/l) (assumed to be reported as mean +/- standard deviation) Monophasic combined high dose group: Baseline= 1.07 +/- 0.2 2 months= 0.94 +/- 0.1 6 months= 0.95 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) Monophasic combined low	equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - yes C2 a. How many participants did not complete treatment in each group? - none C2 b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment) - yes C3 a. For how many participants in each group were no outcome data available? - none C3 b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available) - yes D1 The study had an appropriate length of follow-up - yes D2 The study used a

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				dose group: Baseline= 1.28 +/- 0.2 2 months= 1.58 +/- 0.3 6 months= 1.93 +/- 0.3 No significant difference between baseline and treatment values (p values not reported) Progesterone only group: Baseline= 1.25 +/- 0.1 2 months= 1.66 +/- 0.3 6 months= 1.17 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) Triphasic combined group: Baseline= 1.25 +/- 0.3 2 months= 1.39 +/- 0.4 6 months= 1.12 +/- 0.2 No significant difference between baseline and treatment values (p values not reported) High-density lipoprotein cholesterol (mmol/l) (assumed to be reported as mean +/- standard deviation) Monophasic combined high dose group: Baseline= 1.54 +/- 0.1 2 months= 1.36 +/- 0.1 6 months= 1.33 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) Monophasic combined low dose group:	precise definition of outcome - yes D3 A valid and reliable method was used to determine the outcome - yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - unclear D5 Investigators were kept 'blind' to other important confounding and prognostic factors - unclear Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				Baseline= 1.42 +/- 0.1 2 months= 1.60 +/- 0.1 6 months= 1.52 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) Progesterone only group: Baseline= 1.23 +/- 0.1 2 months= 1.20 +/- 0.1 6 months= 1.30 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) Triphasic combined group: Baseline= 1.51 +/- 0.1 2 months= 1.63 +/- 0.1 6 months= 1.54 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) Low-density lipoprotein cholesterol (mmol/l) (assumed to be reported as mean +/- standard deviation) Monophasic combined high dose group: Baseline= 3.17 +/- 0.4 2 months= 2.99 +/- 0.3 6 months= 3.12 +/- 0.4 No significant difference between baseline and treatment values (p values not reported) Monophasic combined low dose group: Baseline= 3.13 +/- 0.3	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				2 months= 3.35 +/- 0.4 6 months= 3.48 +/- 0.4 No significant difference between baseline and treatment values (p values not reported) Progesterone only group: Baseline= 3.26 +/- 0.2 2 months= 3.46 +/- 0.4 6 months= 3.15 +/- 0.2 No significant difference between baseline and treatment values (p values not reported) Triphasic combined group: Baseline= 3.23 +/- 0.2 2 months= 3.17 +/- 0.3 6 months= 3.35 +/- 0.3 No significant difference between baseline and treatment values (p values not reported) Very low density lipoprotein cholesterol (mmol/l) (assumed to be reported as mean +/- standard deviation) Monophasic combined high dose group: Baseline= 0.49 +/- 0.1 2 months= 0.43 +/- 0.1 6 months= 0.41 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) Monophasic combined low dose group: Baseline= 0.58 +/- 0.1 2 months= 0.72 +/- 0.2	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				6 months= 0.88 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) Progesterone only group: Baseline= 0.57 +/- 0.1 2 months= 0.75 +/- 0.1 6 months= 0.53 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) Triphasic combined group: Baseline= 0.57 +/- 0.1 2 months= 0.63 +/- 0.2 6 months= 0.53 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) High-density lipoprotein cholesterol/total cholesterol (assumed to be reported as mean +/- standard deviation): Monophasic combined high dose group: Baseline= 0.32 +/- 0.1 2 months= 0.30 +/- 0.1 6 months= 0.29 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) Monophasic combined low dose group: Baseline= 0.29 +/- 0.1 2 months= 0.30 +/- 0.1 6 months= 0.27 +/- 0.1	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				No significant difference between baseline and treatment values (p values not reported) Progesterone only group: Baseline= 0.25 +/- 0.1 2 months= 0.23 +/- 0.1 6 months= 0.26 +/- 0.1 No significant difference between baseline and treatment values (p values not reported) Triphasic combined group: Baseline= 0.29 +/- 0.1 2 months= 0.31 +/- 0.1 6 months= 0.29 +/- 0.1 No significant difference between baseline and treatment values (p values not reported)	

A.2 Blood Glucose targets in the pre-conception period

No evidence table

A.3 What is the target value for HbA_{1c} in women with type 1 or type 2 diabetes who are planning pregnancy?

Study details	Participants	Methods	Results	Comments
Full citation Bell,R., Glinianaia,S.V.,	Population All singleton pregnancies to	Methods Linked register analysis of data from	Main outcomes Odds of	Limitations NICE checklist for cohort studies,

Study details	Participants	Methods	Results	Comments
Tennant,P.W.G., Bilous,R.W., Rankin,J., Peri-conception hyperglycaemia and nephropathy are associated with risk of congenital anomaly in women with pre-existing diabetes: A population-based cohort study, Diabetologia, 55, 936-947, 2012 Ref Id 236462 Design Retrospective cohort study Country/ies where the study was carried out United Kingdom Aim of study To determine the risk of major congenital abnormalities during pregnancy in women with type 1 and type 2 diabetes and to determine the effect of clinical and socio-demographic factors risk factors in addition to peri-conception HbA_{1c}. Study dates 1996 to 2008 Funding Study funded by Diabetes UK. Northern Diabetes in Pregnancy Survey (NorDIP) funded by the UK Department of Health/Healthcare Quality Improvement Partnership. Northern Congenital Abnormality Survey (NorCAS) funded by the four Primary Care Trusts in North East England.	women resident in the area captured by NorDIP between 1996 and 2008 and diagnosed with diabetes at least 6 months prior to conception. Sample size N = 1677 Sample size by HbA_{1c} level unknown. Interventions No specific intervention Baseline characteristics Data and p-values not reported with respect to HbA_{1c} levels. Median maternal age at delivery, years (IQR) Type 1: 29 (24 to 33) Type 2: 33 (29 to 37) Median duration of diabetes, years (IQR) Type 1: 2 (6 to 18) Type 2: 2 (1 to 4) Median BMI at baseline, kg/m2 (IQR) Type 1: 25.5 (23 to 29)	NorDIP and NorCAS. Information regarding pre-pregnancy and antenatal HbA_{1c} in women diagnosed with diabetes at least 6 months prior to conception were collected by NorDIP. The total number of registered singleton live and stillbirths was obtained from the UK Office for National Statistics. Data on congenital abnormalities were obtained from NorCAS which reports abnormalities up to age 12, with a maximum of 6 abnormalities per case. This includes those in foetal loss or termination. 23 women (18%) had terminations due to the presence of fetal anomalies. NorCAS uses multiple data sources. Both NorDIP and NorCAS are held on a single linked database. Congenital malformations were coded according to ICD10 codes and categorised using European Surveillance of Congenital Abnormalities (EUROCAT). HbA_{1c} values were DCCT-aligned. Statistical analyses Prevalence rates of congenital abnormalities were compared using relative risks (RR). 95% CIs were calculated using exact methods.	congenital malformations OR per unit increase (percentage) in HbA_{1c} = 1.3 (95% CI 1.2 to 1.4) At a threshold of 6.3% for HbA_{1c} the OR = 5.22 (95% CI 3.15 to 8.32)* for pregnancies being affected by a congenital abnormality LOWESS regression suggested that the risk of pregnancies being affected by a congenital abnormality increased in an approximately linear fashion after the threshold of 6.3%. *Calculated by the NCC-WCH technical team by raising the OR per unit increase to a power of 6.3. Types of	taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. Unclear. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Unclear. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes. C2:

Study details	Participants	Methods	Results	Comments
	Type 2: 34.6 (29 to 40) Inclusion criteria Diagnosis of diabetes at least 6 months prior to conception Singleton pregnancies Live births, still births, late foetal losses or terminations following diagnosis of an anomaly Exclusion criteria Women with gestational diabetes	Independent associations between maternal and neonatal characteristics and congenital abnormalities were assessed using odds ratios (OR) from backward stepwise logistic regression. HbA_{1c} was assessed as a periconception variable using the measurement closest to conception either within three months for 48.4% of women or using mean first trimester value in all other women (up to 14 weeks' gestation). The association between HbA_{1c} as a continuous variable and risk of congenital abnormality was determined using locally weighted scatter plot smoothing.	congenital abnormality, n Nervous system = 16 Eye = 2 Cardiovascular system = 44 Orofacial clefts = 1 Digestive system = 10 Urinary = 12 Genital = 2 Limb = 2 Musculoskeletal = 3 Syndrome (monogenic or unknown) = 11 Multiple anomalies = 9	a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion. N/A C3: a. For how many participants in each group were no outcome data available? Unclear. b. The groups were comparable with respect to the availability of outcome data. Unclear. D. Detection bias D1: The study had an appropriate length of follow-up. Yes. D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome. Yes. D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A

Study details	Participants	Methods	Results	Comments
				Other information None.
Full citation Jensen,D.M., Korsholm,L., Ovesen,P., Beck-Nielsen,H., Moelsted-Pedersen,L., Westergaard,J.G., Moeller,M., Damm,P., Peri-conceptional A_{1C} and risk of serious adverse pregnancy outcome in 933 women with type 1 diabetes, Diabetes Care, 32, 1046-1048, 2009 Ref Id 248370 Design Retrospective cohort Country/ies where the study was carried out Denmark Aim of study To determine whether there is a threshold for peri-conception HbA_{1c} that corresponds to a reduced risk of congenital malformations and perinatal mortality. Study dates 1993 to 1999 Funding The Danish Diabetes Association.	Population Pregnant women with type 1 diabetes. Sample size N = 1215 After excluding multiple and recurrent pregnancies N for analysis = 933 By HbA_{1c} level: < 6.9%: n = 284 ≥ 6.9%: n = 649 Interventions No specific intervention Baseline characteristics P-values were not reported. Mean age, years ± SD 28.6 ± 4.8 Mean BMI, kg/m2 ± SD 23.6 ± 3.5 Mean duration of diabetes, years ± SD 12.3 ± 7.9 Ethnicity	Methods Registry data from the Danish Diabetes Association between 1991 and 1999 were analysed. Data were from eight centres with 75 to 93% coverage. Background population data from 70 089 deliveries recorded by the Danish Health Board in 1995 were used as a comparator group. Perinatal mortality was defined as intrauterine at > 24 weeks' gestation or death during the first 7 days of life Congenital malformations were defined as major if they resulted in death, caused a significant future handicap or required major surgery; all others were classified as minor. Types of congenital malformation were not reported. Alignment with DCCT values for HbA_{1c} was not reported. Statistical analyses Percentages or relative risks (RR) were used to report associations. X2 tests were used to compare outcomes at different levels of HbA_{1c}.	Main outcomes Perinatal mortality, n/N < 6.9%: 6/284 ≥ 6.9%: 25/649 RR = 1.82 (95% CI 0.75 to 4.39)* Congenital malformations < 6.9%: 11/284 ≥ 6.9%: 34/649 RR = 1.35 (95% CI 0.69 to 2.63)* An increased risk of congenital malformations was observed in comparison to a background population of women without diabetes when HbA_{1c} levels were greater than or equal to 10.4% (RR = 3.9, 95% CI: 1.8 to 7.8). *Calculated by the NCC-WCH technical team. Categories of HbA_{1c} were dichotomised at	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. No. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Unclear. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for

Study details	Participants	Methods	Results	Comments
	All women were Caucasian Inclusion criteria Delivery completed after 24 weeks' gestation, or Termination before 24 weeks' gestation because of ultrasound-verified malformations Exclusion criteria Multiple and recurrent pregnancies		6.9% based on the authors' inference that this was the cut-off for increased risk in their categorical analysis.	an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes. C2: a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion. N/A C3: a. For how many participants in each group were no outcome data available? Unclear. b. The groups were comparable with respect to the availability of outcome data. Unclear. D. Detection bias D1: The study had an appropriate length of follow-up. Yes. D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome. Yes. D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A

Study details	Participants	Methods	Results	Comments
				. Other information Comparator group data were not used.
Full citation Miller,E., Hare,J.W., Cloherty,J.P., Dunn,P.J., Gleason,R.E., Soeldner,J.S., Kitzmiller,J.L., Elevated maternal hemoglobin A_{1c} in early pregnancy and major congenital anomalies in infants of diabetic mothers, New England Journal of MedicineN.Engl.J.Med., 304, 1331-1334, 1981 Ref Id 261448 Study design Retrospective review of medical records Country/ies where the study was carried out United States of America Aim of the study To determine whether women with diabetes who deliver infants with congenital malformations had higher HbA_{1c} values in early pregnancy compared with women who did not deliver infants with congenital malformations.	Population Pregnant women with type 1 diabetes. Sample size N = 116 Interventions No specific intervention Characteristics Mean maternal age, years Malformation: 27.2 ± 4.1 No malformation: 27.1 ± 3.5 Male infants, % Malformation: 57.4 No malformation: 53.3 Mean gestational age at HbA_{1c} sampling, weeks Malformation: 9.3 ± 1.8 No malformation: 10.2 ± 2.2 Mean initial maternal HbA_{1c}, %	Methods Medical records were reviewed of all pregnant women with type 1 diabetes who attended prenatal clinics at the Joslin Diabetes Center and Boston Hospital for Women during the study period to determine which women had HBA_{1c} measured at the first clinic visit before 14 weeks' gestation. Gestational age was determined based on the date of the last menstrual period, ultrasound at 16 to 20 weeks and physical examination of the newborn infant. Diabetes was classified using White's classification. HbA_{1c} was measured using HPLC and included the last reversible measurement. Major congenital abnormalities were defined as one causing death or serious handicap or one requiring surgery. Cardiac diagnoses were confirmed by cardiac catheterisation, echocardiography or autopsy.	Main outcomes Malformations, n/N ≤ 8.5%: 2/58 > 8.5%: 13/58 RR = 0.15 (95% CI 0.04 to 0.64)* Types of congenital abnormality, n# Central nervous system = 4 Cardiac = 9 Urinary = 4 Respiratory = 3 Gastrointestinal = 1 Other = 2 *Calculated by the NCC-WCH technical team. #Congenital abnormalities were described	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. Unclear A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. N/A

Study details	Participants	Methods	Results	Comments
Study dates April 1977 to April 1980. Source of funding Grants from the National Institutes of Health, the Diabetes Research and Training Center and the Ames and Biodynamics Corporations.	Malformation: 8.4 ± 1.6 No malformation: 9.5 ± 1.0 White's classification, n Class B: 38 Class C: 32 Class D: 9 Class D4 (benign retinopathy): 26 Class F: 5 Class R: 6 Inclusion criteria Requirement for insulin Initial HbA_{1c} measurement taken before 14 weeks' gestation Delivered at the Boston Hospital for Women Infants examined by one of the authors of the study/their associates at birth Telephone contact with the parents or the infant's paediatrician between 3 and 16 months after birth to determine any anomalies not detected at birth/confirm a final diagnosis of anomalies Exclusion criteria Not reported.	Statistical analyses Mean initial HbA_{1c} was compared between groups using unpaired Student's t-tests.	for each individual infant and diagnoses were not reported according to the main abnormality therefore the total number reported is greater than the number of infants (n = 15) who were diagnosed with any abnormality.	B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes C2: a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). N/A C3: a. For how many participants in each group were no outcome data available? Not reported b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate

Study details	Participants	Methods	Results	Comments
				length of follow-up. Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information HbA_{1c} was measured before 14 weeks' gestation. The mean gestational age and standard deviation for each group suggested that HbA_{1c} was measured at or before 12 weeks in most women. Because HbA_{1c} measurement provides an average of glycaemic control for the preceding 3 months this study was included as measuring HbA_{1c} pre-pregnancy (or peri-conception) and rated down for indirectness accordingly.
Full citation Miodovnik,M., Skillman,C., Holroyde,J.C., Butler,J.B., Wendel,J.S., Siddiqi,T.A., Elevated maternal glycohemoglobin in early pregnancy and spontaneous	Population Pregnant women with type 1 diabetes. Sample size	Methods The study group consisted of 116 pregnancies in 75 women. At enrolment medical and obstetric histories were taken. Pregnancy dating was based on menstrual history as	Main outcomes Spontaneous miscarriage in relation to HbA1 measured at study entry, n/N	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to

Study details	Participants	Methods	Results	Comments
abortion among insulin-dependent diabetic women, American Journal of Obstetrics and GynecologyAm.J.Obstet.Gynecol., 153, 439-442, 1985 Ref Id 261434 Design Prospective cohort Country/ies where the study was carried out United States of America Aim of the study To determine whether poor glucose control affected the incidence of spontaneous abortions in pregnant women with type 1 diabetes. Study dates 1978 to 1984 Funding Part funded by grants from the National Institutes of Health, Diabetes in Pregnancy, United States Public Health Service Training in Perinatal Care and Research and the National Institutes of Health Clinical Research Center.	N = 75 (116 pregnancies) Interventions No specific intervention. Baseline characteristics Mean maternal age, years ± SD HbA_{1c} measured by column chromatography Term delivery = 25.7 ±0.5 Spontaneous abortion = 23.6 ± 1.0 HbA_{1c} measured by HPLC Term delivery = 24.1 ± 0.9 Spontaneous abortion = 24.2 ± 1.1 Mean duration of diabetes, years ± SD HbA_{1c} measured by column chromatography Term delivery = 10.9 ± 0.7 Spontaneous abortion = 12.4 ± 1.4 HbA_{1c} measured by HPLC Term delivery = 9.4 ± 1.3 Spontaneous abortion = 11.8 ± 1.1 Mean gestational age when HbA1 measured, weeks ± SD	well as physical and ultrasound examinations. Women were seen every 1 to 2 weeks throughout pregnancy. The goal of treatment for all women was to obtain a fasting blood glucose < 100mg/dl (5.6mmol/l) and a 1.5 hour post-prandial blood glucose < 140mg/dl (7.8mmol/l). Glycaemic control was obtained using split-dose regimen of insulin and diet regulation. Insulin therapy included both short- and intermediate-acting insulin. HbA1 was measured using HPLC in women who delivered between 1978 and 1980 and using column chromatography in women who delivered between 1980 and 1984. HbA1 was measured at entry and once during each trimester. A threshold of 12.0% for HbA1 was applied post-hoc. Spontaneous miscarriages were defined as those occurring between 5 and 15 weeks' gestation. Women with pregnancies which continued past 20 weeks' gestation delivered between 28 and 42 weeks' gestation. Statistical analyses Two different laboratory techniques were used to measure HbA_{1c} therefore women were grouped separately in analyses.	< 12%† = 14/89 ≥ 12%† = 12/27 RR = 0.35 (95% CI: 0.18 to 0.66)* *Calculated by the NCC-WCH technical team. Data for all women were analysed together, regardless of how HbA1 was measured. †An HbA1 of 12.0% corresponds to an HbA_{1c} of 10.9%.	treatment groups was unrelated to potential confounding factors. Unclear A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. No A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? N/A

Study details	Participants	Methods	Results	Comments
	HbA_{1c} measured by column chromatography Term delivery = 8.9 ± 0.2 Spontaneous abortion = 8.1 ± 0.5 HbA_{1c} measured by HPLC Term delivery = 9.1 ± 0.3 Spontaneous abortion = 8.5 ± 0.5 Inclusion criteria Not reported. Exclusion criteria Pregnancies which resulted in congenital malformations	Categorical variables were analysed using either X2 tests or Fisher's exact tests.		b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). N/A C3: a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A - threshold applied post hoc D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A - threshold

Study details	Participants	Methods	Results	Comments
				applied post hoc Other information A value of 12% for HbA1 corresponds to an HbA_{1c} of 8.8% using a standard conversion formula.
Full citation Diabetes and Pregnancy Group,France, French multicentric survey of outcome of pregnancy in women with pregestational diabetes, Diabetes Care, 26, 2990-2993, 2003 Ref Id 261443 Design Cross sectional Country/ies where the study was carried out France Aim of study To assess whether pregnancy outcomes in women with diabetes had improved ten years after the definition of the St Vincent's targets to reduce morbidity in this population. Study dates January 2000 to December 2001 Funding Not reported.	Population All women with type 1 or type 2 diabetes and a single pregnancy who delivered between January 2000 and December 2001. Sample size N = 435 By HbA_{1c} level: ≤ 8.0%: n = 315 > 8.0%: n = 120 Interventions No specific intervention. Baseline characteristics Data were not reported by HbA_{1c} levels. P-values not reported. Diabetes type, n/N Type 1: 289/435 (66%) Type 2: 146/435 (34%) First trimester HbA_{1c} > 8.0%, n/n	Methods Twelve tertiary perinatal centres participated in the study. All data were prospectively collected using the Obstetrical Quality Indicators and Data Collection aggregated database including: Preconception care HbA_{1c} > 8.0% during the first and third trimesters Retinopathy Nephropathy Gestational hypertension or pre-eclampsia Pregnancy outcomes (perinatal mortality, major congeital malformations, pre-term delivery) Macrosomia Mode of delivery Neonatal complications Preconception care included information on optimising glycaemic control before pregnancy and assessment of complications, diet, intensification of self-monitoring of blood glucose and optimisation of insulin.	Main outcomes Perinatal mortality, n/N ≤ 8.0%: 8/315 > 8.0%: 11/120 RR = 0.28 (95% CI 0.11 to 0.68)* Congenital malformations, n/N ≤ 8.0%: 8/315 > 8.0%: 10/120 RR = 0.30 (95% CI 0.12 to 0.74)* *Calculated by the NCC-WCH technical team.	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. No. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear - very little demographic data presented. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Unclear. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A

Study details	Participants	Methods	Results	Comments
	Type 1: 88/289 (30%) Type 2: 32/146 (22%) Data for maternal age, BMI, ethnicity were not reported. Inclusion criteria Women with pre-existing type 1 or type 2 diabetes Singleton pregnancies Delivery between January 2000 and December 2001 Exclusion criteria Women with gestational diabetes Women with multiple pregnancies	HbA_{1c} was obtained in the first trimester. Actual values were not available for HbA_{1c} < 8.0% therefore optimal pre-pregnancy control was assumed to be ≤ 8.0%. Alignment with DCCT values for HbA_{1c} was not reported. Foetal death was defined as ≥ 22 weeks' gestation or > 500g in weight. Neonatal mortality was defined as before the 28th day of life. Major congenital malformations were classified according to EUROCAT. Types of congenital malformation were not reported. Four terminations were performed due to the presence of major congenital abnormalities. Statistical analyses Group comparisons were performed using either X2 tests or Fisher's exact tests where appropriate. Logistic regression was used to assess independent effects of variables on pregnancy outcomes. Results were presented as odds ratios with 95% CIs.		B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes. C2: a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion. N/A C3: a. For how many participants in each group were no outcome data available? Unclear. b. The groups were comparable with respect to the availability of outcome data. Unclear. D. Detection bias D1: The study had an appropriate length of follow-up. Yes. D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome. Yes.

Study details	Participants	Methods	Results	Comments
				D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A Other information None.
Full citation Suhonen,L., Hiilesmaa,V., Teramo,K., Glycaemic control during early pregnancy and fetal malformations in women with type I diabetes mellitus, Diabetologia, 43, 79-82, 2000 Ref Id 261445 Design Retrospective data analysis Country/ies where the study was carried out Finland Aim of study To assess the risk of fetal malformations in women with type 1 diabetes compared to a background population and to relate this risk to glycaemic control during early pregnancy. Study dates 1988 to 1997 Funding Not reported.	Population Cases Pregnant women with type 1 diabetes and their offspring who attended the Department of Obstetrics and Gynaecology at Helsinki University Central Hospital between 1988 and 1997. Controls Offspring from consecutive pregnancies in unselected residents of the city of Kerava who attended routine screening at 16 to 19 weeks' gestation in 1993 and 1994. Data for controls were not used in NCC-WCH analyses. Sample size Cases N = 691 pregnancies Offspring = 709 (16 sets of twins, one set of triplets)	Methods Cases were typically registered at the hospital between 5 and 10 weeks' gestation. In 93% the first visit was < 14 weeks' gestation. Infants were examined for malformations between 2 and 5 days after birth. Outcomes of pregnancies were ascertained from medical records for both mothers and infants. Congenital malformations were defined as major if fatal, likely to cause serious handicap or required surgery; all others were classed as minor. Five women had terminations due to the presence of congenital abnormalities. Alignment with DCCT values for HbA_{1c} was not reported. HbA_{1c} values were compared with Finnish norms. Statistical analyses	Main outcomes Congenital malformations, n/N < 5.6%: 1/47 ≥ 5.6%: 25/616 RR = 0.50 (95% CI 0.07 to 3.61)* *Calculated by the NCC-WCH technical team. Categories of HbA_{1c} were dichotomised at 5.6% by the NCC-WCH technical team based on the cut-off for normal values quoted in the study.	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. No. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Unclear. B2: Participants receiving care were

Study details	Participants	Methods	Results	Comments
	By HbA_{1c} levels: < 5.6%: n = 47 ≥ 5.6%: n = 616 Controls N = 729 pregnancies Offspring = 735 (6 sets of twins) Interventions No specific intervention. Baseline characteristics P-values were not reported. Overall mean duration of diabetes, years ± SD 14.5 ± 7.9 Ethnicity 98% of diabetic women and controls were Caucasian. No information regarding mean maternal age, parity or BMI was reported in relation to HbA_{1c} levels. Inclusion criteria Cases: Pregnant women with type 1 diabetes	Power calculations suggested a required sample size of 602 per group for a 4% vs. 8% malformation rate with 90% power and nominal p-value = 0.05. Continuous variables were analysed using Student's t-tests or Mann-Whitney U tests. Proportions were compared using rate difference and 95% CI. Relative risks and 95% CIs were calculated for malformations for different values of HbA_{1c}.		kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear. C2: a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion. N/A C3: a. For how many participants in each group were no outcome data available? Unclear. b. The groups were comparable with respect to the availability of outcome data. Unclear. D. Detection bias D1: The study had an appropriate length of follow-up. Yes. D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome.

Study details	Participants	Methods	Results	Comments
	Controls: Pregnant women residing within Kerava Attended ultrasound screening between 16 and 19 weeks' gestation Exclusion criteria Cases: None described Controls: Type 2 diabetes Requirement of insulin during pregnancy			Unclear. D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A Other information None.
Full citation Tennant,P.W., Glinianaia,S.V., Bilous,R.W., Rankin,J., Bell,R., Pre-existing diabetes, maternal glycated haemoglobin, and the risks of fetal and infant death: a population-based study, Diabetologia, 57, 285-294, 2014 Ref Id 305877 Study design Retrospective cohort study Country/ies where the study was carried out UK Aim of the study To investigate the association between pre-existing diabetes and the risks of fetal and infant death in normally formed offspring, and to quantify the contribution of	Population All singleton pregnancies to women resident in the area captured by NorDIP between 1996 and 2008 and diagnosed with diabetes at least 6 months prior to conception. Sample size N=1548 Sample size by HbA_{1c} level unknown. Interventions No specific intervention. Characteristics Median maternal age at delivery, years (IQR)	Methods The total number of singleton live births and fetal and infant deaths were obtained from the UK Office for National Statistics and the Northern Perinatal Morbidity and Mortality Survey (PMMS), respectively. The number of normally formed offspring was determined by subtracting the number of NorCAS registrations. Mode of birth not reported. 'Late miscarriages' are the spontaneous loss of a fetus at 20 to 30 completed weeks gestation. 'Stillbirths' are deliveries of a fetus showing no signs of life at 24 or more completed weeks of gestation.	Main outcomes Odds of fetal and infant death Increasing HbA_{1c} concentration above values of 49mmol/mol (6.6%) increase the odds of fetal and infant death Adjusted OR = 1.02 (95% CI 1.00 to 1.04) P=0.04 Types of fetal or infant death, n Fetal death = 46 Late miscarriage = 5 Still birth = 41 (antepartrum	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. Unclear. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear. B. Performance bias B1: The comparison groups received

Study details	Participants	Methods	Results	Comments
glycaemic control. Study dates 1 Januray 1996 to 31 December 2008 Source of funding The study was part funded by Diabetes UK. The NorDIP, PMMS and NorCAS are funded by Public Health England.	30 (25 to 34) Median periconceptional HbA_{1c} concentrations, mmol/mol (IQR) 62 (51 to 76) Median third trimester HbA_{1c} concentrations, mmol/mol (IQR) 50 (43 to 58) Median BMI at baseline, kg/m2 (IQR) 27 (24 to 32) Ethnicity and smoking not reported. Inclusion criteria Singleton pregnancies Pre-existing diabetes (type 1 or type 2) at least 6 months before conception Delivered at or after 20 completed weeks of gestation Exclusion criteria Women with gestational diabetes Pregnancies identified from the Northern Congenital Abnormality Survey (NorCAS) complicated by major congenital anomalies, which have previously been shown to be asssociated with both	'Late stillbirths' are stillbirths at 28 or more completed weeks of gestation. 'Antepartrum stillbirths' are stillbirths where the fetus dies before the onset of labour. 'Neonatal deaths' are deaths, after live birth, within the first 28 days of life. 'Postnatal deaths' are deaths, after live birth, of an infant aged 28 days or more, but less than one year. 'Infant deaths' comprise neonatal deaths and postnatal deaths. Statistical analyses Periconception HbA_{1c} concentration was chosen as a reasonable surrogate of preconception HbA_{1c} concentration as first-trimester HbA_{1c} correlated highly with preconception HbA_{1c}. Prevalence rates of fetal or infant deaths were compared using relative risks (RR), 95% CIs were calculated using exact method. Odds ratios (ORs) and 95% CIs for all variables with hypothesised influences on fetal and/or infant death were analysed in relation to fetal death, late still birth, infant death, fetal and infant death combined, and late still birth and infant death combined within a series	stillbirth = 38, intrapartum stillbirth = 3) Infant death = 10 Neonatal death = 6 Postnatal death = 4	the same care apart from the intervention(s) studied. Unclear. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes. C2: a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion. N/A C3: a. For how many participants in each group were no outcome data available? Unclear. b. The groups were comparable with respect to the availability of outcome data. Unclear. D. Detection bias D1: The study had an appropriate length of follow-up. Yes.

Study details	Participants	Methods	Results	Comments
	pre-existing diabetes and the risk of fetal and infant death	of logit-linked generalised estimating equations. Adjusted ORs were estimated from backward stepwise logistic regression. HbA_{1c} was assessed as a periconception variable using the measurement closest to conception either within three months for 48.4% of women or using mean first trimester value in all other women (up to 14 weeks' gestation). Third trimester HbA_{1c} was examined only in relation to deliveries at >28 weeks of gestation. The association between HbA_{1c} as a continuous variable and risk of fetal and infant death was determined using locally weighted scatter plot smoothing.		D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome. Yes. D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A Other information
Full citation Greene,M.F., Hare,J.W., Cloherty,J.P., Benacerraf,B.R., Soeldner,J.S., First-trimester hemoglobin A1 and risk for major malformation and spontaneous abortion in diabetic pregnancy.[see comment], Teratology, 39, 225-231, 1989 Ref Id 261456 Design Retrospective cohort Country/ies where the study was	Population Women with type 1 diabetes presenting at the Joslin Diabetes Centre prenatal clinic in the study period. Sample size N = 303 (no explanation was provided for missing data for n = 31 women) By HbA1 level: ≤ 9.3%: n = 99	Methods All eligible patients within the study period were included. HbA1 was measured rather than HbA_{1c} HbA1 values were therefore not DCCT-aligned. When more than HbA1 measurement was available in medical records the earliest recorded value was used. Severity of diabetes was classified according to White's criteria.	Main outcomes Congenital malformations, n/N ≤ 9.3%: 3/99† > 9.3%: 17/151#† RR = 0.27 (95% CI 0.08 to 0.90)* #One woman was excluded from analyses due an elected termination.	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. No.

Study details	Participants	Methods	Results	Comments
carried out United States of America Aim of study To examine the relationship between metabolic control in pregnant women with diabetes and congenital malformations. Study dates December 1983 to December 1987 Funding Not reported.	> 9.3%: n = 152 Interventions No specific intervention. Baseline characteristics Data were not reported according to HbA_{1c} level. Mean maternal age, years ± SD No major malformation: 29.2 ± 4.7 Spontaneous abortion: 29.5 ± 5.3 Major malformation: 27.1 ± 3.9 P-value not significant Mean duration of diabetes before pregnancy, years ± SD No major malformation: 13.4 ± 7.0 Spontaneous abortion: 12.0 ± 7.7 Major malformation: 8.9 ± 5.1 P-value = 0.025 Inclusion criteria Patients presenting within the study dates with a known outcome	Routine ultrasound was undertaken at 16 to 19 weeks' gestation. The attending examiner was not aware of the first trimester HbA1 value. Diagnosis of spontaneous abortion was made using serial ultrasound. Six paediatricians performed all of the neonatal examinations. Spontaneous abortion was defined as an empty intrauterine gestational sac, ultrasonographic identification of a foetus without cardiac motion or histological identification of a trophoblast. Congenital malformations were major if fatal, required surgery to correct or were of major anatomical/cosmetic concern. Five women had terminations due to the presence of congenital abnormalities. Three women whose foetuses were diagnosed with an abnormality in the second trimester did not have a termination. These three pregnancies resulted in the fatality of the infant during or after birth. Statistical analyses Analysis of continuous variables was carried out using ANOVA. Risk ratios were calculated using the Mantel-Haenszel X2 test.	*Calculated by the NCC-WCH technical team. Categories were dichotomised for analysis at 9.3% based on the use of this mean HbA1 value being used as the cut-off for the referent group by the study authors. †HbA1 was converted to HbA_{1c} using a standard formula by the NCC-WCH technical team. An HbA1 of 9.3% corresponds to an HbA_{1c} of 8.4%. One case of each of the following abnormalities was observed: Tetralogy of Fallot Diaphragmatic hernia Atrioventricular canal hydrops fatalis Bilateral renal agenesis oligohydramnios	A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Unclear. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear. C2: a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion. N/A C3: a. For how many participants in each group were no outcome data available? Not reported: no explanation was provided for missing data for 31 women.

Study details	Participants	Methods	Results	Comments
	≤ 12 weeks' gestation Exclusion criteria A total of 21 patients were excluded: 2 suffered first trimester spontaneous abortions 9 transferred their care to other physicians 10 were lost to follow-up One additional woman was excluded from analyses due to a termination.	P-values < 0.05 were taken to be significant.	Bilateral renal hypoplasia Three cases of anencephaly were observed.	b. The groups were comparable with respect to the availability of outcome data. Unclear. D. Detection bias D1: The study had an appropriate length of follow-up. Yes. D2: The study used a precise definition of outcome. Yes for spontaneous abortion, unclear for classification of major congenital abnormalities. D3: A valid and reliable method was used to determine the outcome. Yes for spontaneous abortion, unclear for congenital malformations. D4: Investigators were kept 'blind' to participants' exposure to the intervention. Yes for spontaneous abortion - examiner did not know first trimester HbA_{1c} status. D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A Other information Risks of congenital malformations were presented by categories of HbA_{1c}.

A.4 Ketone monitoring in the pre-conception period

No evidence table

A.5 What is the effectiveness of specialist teams for pregnant women with diabetes compared to separate obstetric and endocrinology teams?

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation Owens,L.A., Avalos,G., Kirwan,B., Carmody,L., Dunne,F.P., Changing clinical practice can improve clinical outcomes for women with pre-gestational diabetes mellitus, Irish Medical Journal, 105, 9-11, 2012 Ref Id 224407 Country/ies where the study was carried out Ireland Study type Prospective observational study Aim of the study To compare pregnancy outcomes before and after the introduction of a dedicated combined antenatal/diabetes	Sample size 272 pregnancies (number of women not reported) Characteristics Type 1 diabetes: Multidisciplinary team= 87 (52%) Non-multidisciplinary team= 80 (77%) p value not reported Type 2 diabetes: Multidisciplinary team= 81 (48%) Non-multidisciplinary team= 24 (23%) p value not reported Pre-pregnancy care: Multidisciplinary team= 52%	Interventions Multidisciplinary team (n= 168) Non-multidisciplinary team (n= 104)	Details Two cohorts of women were included - one from 2005-7 and another from 2008-10. Details of the care received by the women in 2005-7 is not reported (assumed to be a non-multidisciplinary team). The women who were pregnant in 2008-10 received care from a dedicated combined antenatal/diabetes clinic and pre-pregnancy care clinic, delivered by specialist diabetes and obstetric staff (multidisciplinary team). Locally developed clinical care guidelines based on NICE guidance were used. All women were invited and encouraged to attend pre-pregnancy care, which consisted of education, contraception advice, provision of folic acid for 12	Results For all of the following results, apart from perinatal mortality and HbA_{1C}, the paper only reported the percentage of women and not the raw data. The raw data were calculated by the NCC-WCH, and so rounding errors may be present. Caesarean section Multidisciplinary team= 113/168* (67%) Non-multidisciplinary team= 58/104 (56%) p=0.01 OR= 1.63 (95% CI 0.98 to 2.70)** Elective section Multidisciplinary team= 92/168* (55%) Non-multidisciplinary team= 24/104 (23%) p=0.01	Limitations There are some anomalies in the data (see footnotes in 'Results' section). NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation to treatment groups was unrelated to potential confounding factors - Unclear A2 Attempts were made within the design or analysis to balance the comparison groups for potential confounders - No A3 Groups were

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
clinics and pre-pregnancy care clinics delivered by specialist diabetes and obstetric staff. Study dates One group from 2005-2007, one group from 2008-2010 Source of funding None reported	Non-multidisciplinary team= 28% p<0.05 Folic acid (5mg) Multidisciplinary team= 62% Non-multidisciplinary team= 41% p value not significant Achieved target HbA_{1C} at booking of <7% (<53mmol) Multidisciplinary team= 63% Non-multidisciplinary team= 48% p<0.05 Mean BMI at booking (kg/m2) Type 1 diabetes: Multidisciplinary team= 26 +/- 4.81 Non-multidisciplinary team= 26 +/- 4.32 p value not significant Type 2 diabetes Multidisciplinary team= 33 +/- 6.4 Non-multidisciplinary team= 30 +/- 5.6 p value not significant % with BMI > 30 kg/m2 at		weeks, discussion of glycaemic targets, initiation/intensification of insulin therapy, prevention and treatment of hypoglycaemia, discontinuation of teratogenic drugs where appropriate, management of blood pressure and diabetes-related complications. Large for gestational age was defined as birth weight above the 90th centile. P values were reported for the comparison of some outcomes, however, the method of analysis was not reported.	Emergency section Multidisciplinary team= 45/168* (27%) Non-multidisciplinary team= 34/104 (33%) p value not significant HbA_{1C} (mmol) in first trimester Type 1 diabetes Multidisciplinary team= 60 +/- 6 Non-multidisciplinary team= 63 +/- 6 p<0.0001 MD= -3.00 (95% CI -4.47 to -1.53)** HbA_{1C} (mmol) in first trimester Type 2 diabetes Multidisciplinary team= 54 +/- 7 Non-multidisciplinary team= 61 +/- 5 p<0.0001 MD= -7.00 (95% CI -8.43 to -5.57)** HbA_{1C} (mmol) in second trimester Type 1 diabetes Multidisciplinary team= 50 +/- 1.1 Non-multidisciplinary team= 51 +/- 1.2	comparable at baseline, including all major confounding and prognostic factors - Unclear B1 Comparison groups received the same care apart from the intervention(s) studied - No B2 Participants receiving care were kept 'blind' to treatment allocation - N/A B3 Individuals administering care were kept 'blind' to treatment allocation - N/A C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? - N/A C2 b. Groups were comparable for treatment completion - Yes C3 a. For how

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	booking Type 1 diabetes Not reported Type 2 diabetes Multidisciplinary team= 79% Non-multidisciplinary team= 50% p<0.05 Inclusion criteria All women with diabetes for greater than 6 months before the index pregnancy Exclusion criteria None reported			p<0.0001 MD= -1.00 (95% CI -1.28 to -0.72)** HbA_{1C} (mmol) in second trimester Type 2 diabetes Multidisciplinary team= 41 +/- 0.7 Non-multidisciplinary team= 46 +/- 1.0 p<0.0001 MD= -5.00 (95% CI -5.22 to -4.78)** HbA_{1C} (mmol) in third trimester Type 1 diabetes Multidisciplinary team= 46 +/- 0.9 Non-multidisciplinary team= 49 +/- 1.1 p<0.0001 MD= -3.00 (95% CI -3.25 to -2.75)** HbA_{1C} (mmol) in third trimester Type 2 diabetes Multidisciplinary team= 42 +/- 0.6 Non-multidisciplinary team= 41 +/- 0.9 p<0.0001 MD= 1.00 (95% CI 0.80 to 1.20)**	many participants in each group were no outcome data available? - None C3 b. Groups were comparable with respect to the availability of outcome data - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - N/A D5 Investigators were kept 'blind' to other important confounding and prognostic factors - N/A Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				Live birth rate Multidisciplinary team= 155/168 (92%) Non-multidisciplinary team= 77/104 (74%) p<0.0001 Perinatal mortality rate Multidisciplinary team= 1 (0.65%)*** Non-multidisciplinary team= 5 (6.2%)*** p<0.0001 OR= 0.12 (95% CI 0.01 to 1.03)** Miscarriage Multidisciplinary team= 13/168 (8%) Non-multidisciplinary team= 23/104 (22%) p<0.0001 OR= 0.30 (95% CI 0.14 to 0.61)** Still birth Multidisciplinary team= 2/168 (1%) Non-multidisciplinary team= 4/104 (4%) p<0.0001 OR= 0.30 (95% CI 0.05 to 1.67)** Large for gestational age babies	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				Type 1 diabetes Multidisciplinary team= 44/168 (26%)**** Non-multidisciplinary team= 31/104 (30%) p<0.05 OR= 0.84 (95% CI 0.49 to 1.44)** Large for gestational age babies Type 2 diabetes Multidisciplinary team= 42/168 (25%) Non-multidisciplinary team= 18/104 (18%) p value not reported OR= 1.59 (95% CI 0.86 to 2.95)** Neonatal ICU admission Multidisciplinary team= 94/168 (56%) Non-multidisciplinary team= 63/104 (61%) p value not significant OR= 0.83 (95% CI 0.50 to 1.36)** *The raw data were calculated by the NCC-WCH based on the percentages reported in the paper. The number of elective caesarean sections and the number of emergency caesarean sections do not add up to the total number of	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				caesarean sections, as raw data or as the percentages reported in the study. **Calculated by the NCC-WCH based on results reported in the paper ***These are the raw data and percentages as reported in the paper. It is not clear which denominator was used. ****This is reported as 26% in this paper. However, the same authors published a paper on the same study in Diabetes Care (Owens, 2012), which reports this as 16%. It is assumed that 26% is correct, as it reflects the reported p value more accurately (the same p value is reported in both papers).	
Full citation Wilson,N., Ashawesh,K., Kulambil Padinjakara,R.N., Anwar,A., The multidisciplinary diabetes-endocrinology clinic and postprandial blood glucose monitoring in the management of gestational diabetes: impact on maternal and neonatal outcomes, Experimental and Clinical Endocrinology	Sample size 96 women Characteristics Age at booking (years): Multidisciplinary team= 31.40 (+/- 4.85) Non-multidisciplinary team= 29.71 (+/- 6.02) p value not significant Gestation at booking (weeks): Multidisciplinary team=	Interventions Multidisciplinary team (n= 47) Non-multidisciplinary team (n= 49)	Details Two cohorts were randomly selected from hospital held lists (details of randomisation method not provided) of women attending clinics at a hospital. 50 women were selected for each cohort. One cohort was from 2000 to 2002 and the other from 2006 to 2008. From 2003 to 2005, an endocrinology-antenatal care clinic was introduced at the hospital, therefore the cohort of women from 2006 to 2008	Results Vaginal delivery Multidisciplinary team= 22/47 (46.8%) Non-multidisciplinary team= 21/49 (43.8%) p value not reported OR= 1.17 (95% CI 0.52 to 2.62)** Assisted delivery (including forceps and fentouse) Multidisciplinary team= 3/47 (6.4%)	Limitations It is not clear whether the groups were comparable in terms of BMI, as conflicting data were reported in the text (see 'Characteristics' section). NICE guidelines manual. Appendix D: Methodology checklist: Cohort

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
and Diabetes, 117, 486-489, 2009 Ref Id 224567 Country/ies where the study was carried out UK Study type Retrospective observational study Aim of the study To audit the introduction of a multidisciplinary endocrinology-antenatal clinic and diabetes specialist nurse Study dates One cohort from Jan 2000 to Dec 2002, one cohort from Jan 2006 to Feb 2008 Source of funding None reported	11.90 (+/- 2.98) Non-multidisciplinary team= 13.79 (+/- 4.23) p<0.01 BMI≥30kg/m2* Multidisciplinary team= 22 (47.8%) Non-multidisciplinary team= 12 (34.3%) p<0.05 Management at diagnosis with diet/lifestyle modification alone Multidisciplinary team= 43 (91.5%) Non-multidisciplinary team= 45 (91.8%) p value not reported Management at diagnosis with insulin Multidisciplinary team= 4 (8.5%) Non-multidisciplinary team= 4 (8.2%) p value not reported Management at birth with diet/lifestyle advice alone Multidisciplinary team= 9 (19.1%) Non-multidisciplinary team= 32 (65.3%) p value not reported		received care through this clinic (multidisciplinary team). It is not reported how women with diabetes were managed in pregnancy prior to this, including those in the 2000 to 2002 cohort. It is assumed that this cohort of women received non-specialised care (non-multidisciplinary team). The endocrinology-antenatal care clinic included an endocrinologist, obstetrician, diabetes specialist nurse, and dietitian. Patients were issued with a home blood glucose monitor and advised to maintain their 1 hour postprandial blood glucose at 7.8mmol/L or below. Patient information was obtained from clinic-held summaries, obstetric notes, and patient held pregnancy records retained in the hospital after birth. Birthweight centiles were calculated using the ImsGrowth programme obtained from the Child Growth Foundation. Data were compared using X2 and unpaired two-tail t-test as appropriate.	Non-multidisciplinary team= 4/49 (8.3%) p value not reported OR= 0.77 (95% CI 0.16 to 3.63)** Emergency caesarean Multidisciplinary team= 7/47 (14.9%) Non-multidisciplinary team= 9/49 (18.8%) p value not reported Elective caesarean Multidisciplinary team= 15/47 (31.9%) Non-multidisciplinary team= 14/49 (29.2%) p value not reported Any caesarean Multidisciplinary team= 22/47 (47%) Non-multidisciplinary team= 23/49 (47%) p value not reported OR= 1.41 (95% CI 0.92 to 2.17)** HbA_{1C} trimester 1 (mean +/- standard deviation) Multidisciplinary team= 6.144 +/- 0.384 Non-multidisciplinary team= 6.067 +/- 1.139 p value not significant OR= 0.00 (95% CI -0.33 to	studies A1 Method of allocation to treatment groups was unrelated to potential confounding factors - Unclear A2 Attempts were made within the design or analysis to balance the comparison groups for potential confounders - No A3 Groups were comparable at baseline, including all major confounding and prognostic factors - Unclear B1 Comparison groups received the same care apart from the intervention(s) studied - Yes B2 Participants receiving care were kept 'blind' to treatment allocation - N/A B3 Individuals administering care were kept 'blind' to treatment allocation - N/A C1 All groups were followed up for an

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Management at birth with insulin Multidisciplinary team= 38 (80.9%) Non-multidisciplinary team= 17 (34.7%) p value not reported Ethnicity: White Multidisciplinary team= 42.6% Non-multidisciplinary team= 51.0% p value not significant South Asian (including Indians, Pakistanis, Bangladeshis, and other Asians) Multidisciplinary team= 38.2% Non-multidisciplinary team= 34.6% P value not significant *These are the data as reported in a table in the study. In the text, however, it states 'In [the multidisciplinary team cohort] only one patient had their BMI recorded compared to 14 in [the non-multidisciplinary team cohort]'. Inclusion criteria			0.33)** HbA_{1C} trimester 2 (mean +/- standard deviation) Multidisciplinary team= 5.737 +/- 0.527 Non-multidisciplinary team= 5.911 +/- 1.184 p value not significant OR= -0.20 (95% CI -0.57 to 0.17)** HbA_{1C} trimester 3 (mean +/- standard deviation) Multidisciplinary team= 5.855 +/- 0.579 Non-multidisciplinary team= 6.288 +/- 0.934 p<0.001 OR= -0.40 (95% CI -0.70 to -0.10)** Birthweight (g, mean +/- standard deviation) Multidisciplinary team= 3269 +/- 675 Non-multidisciplinary team= 3567 +/- 700 p<0.05 Birthweight centile (mean +/- standard deviation) Multidisciplinary team= 57.01 +/- 31.18 Non-multidisciplinary team= 72.47 +/- 29.56 p<0.05	equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? - N/A C2 b. Groups were comparable for treatment completion - Yes C3 a. For how many participants in each group were no outcome data available? - None C3 b. Groups were comparable with respect to the availability of outcome data - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants'

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Not reported Exclusion criteria Incomplete patient notes and/or missing data			Admission to SCBU Multidisciplinary team= 5/47 (10.6%) Non-multidisciplinary team= 16/49 (32.7%) p<0.01 OR= 0.25 (95% CI 0.08 to 0.74)** The infants admitted to SCBU in the multidisciplinary team cohort 'remained in hospital significantly longer' than in the non-multidisciplinary team cohort (p < 0.05, actual data not reported) **Calculated by the NCC-WCH based on results reported in the paper	exposure to the intervention - N/A D5 Investigators were kept 'blind' to other important confounding and prognostic factors - N/A Other information Of the women randomly selected for the audit, 3 women from the multidisciplinary team cohort and 1 woman from the non-multidisciplinary team cohort were excluded from the analyses as their data were incomplete
Full citation Dunne,F.P., Avalos,G., Durkan,M., Mitchell,Y., Gallacher,T., Keenan,M., Hogan,M., Carmody,L.A., Gaffney,G., TLANTIC,D.I.P., ATLANTIC DIP: pregnancy outcome for women with pregestational diabetes along the Irish Atlantic seaboard, Diabetes Care, 32, 1205-1206,	Sample size 104 pregnancies (84 women*) * indicates information or data reported in Dunne (2012) 'ATLANTIC DIP: Pregnancy outcomes for women with type 1 and type 2 diabetes' which reported on the same cohort of women Characteristics	Interventions Centralised care (n= 31) Peripheral care (n= 73)	Details Women were managed according to local guidelines. Values of HbA_{1c} were taken at the first visit, then at 12, 24, and 36 weeks, and before delivery. Large for gestational age was defined as birth weight greater than 4kg. Although significant differences were reported, the method of analysis was not reported.	Results Live births: Central= 25/31 (81%) Peripheral= 54/73 (74%) p value not reported Miscarriage: Central= 6/31 (19%) Peripheral= 17/73 (23%) p value not reported OR= 0.79 (95% CI 0.28 to 2.24)**	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation to treatment groups was unrelated to potential confounding factors - Unclear A2 Attempts were

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
2009 Ref Id 224395 Country/ies where the study was carried out Ireland Study type Prospective observational study Aim of the study To outline pregnancy outcomes in women with type 1 and type 2 diabetes Study dates 2006 to 2007 (months not given) Source of funding 'The costs of publication of the article were defrayed in part by the payment of page charges. The article must therefore be hereby marked 'advertisement' in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.'	The characteristics reported below were not reported separately for women who attended a peripheral or a central hospital Type of diabetes: Type 1= 80/104 (77%) Type 2= 24/104 (23%) Mean age at delivery* (SD): Type 1 diabetes= 33 +/- 5.7 years Type 2 diabetes= 36 +/- 4.4 years p=0.04 Duration of diabetes (years): Type 1 diabetes= 14 years Type 2 diabetes= 5 years p=0.0001 Complications at booking: Retinopathy= 16 (18%) Renal disease= 7 (8%) Hypertension= 3 (3%) All of these complications were in women with type 1 diabetes* Body Mass Index (BMI): BMI < 25 kg/m2 = 32%* BMI > 25 kg/m2 to < 30 kg/m2 = 50%			Stillbirth: Central= 0/31 (0%) Peripheral= 2/73 (3.6%) p value not reported OR= 0.45 (95% CI 0.02 to 9.73)** Small for gestational age: Central= 0/31 (0%) Peripheral= 4/73 (7%) p value not reported Large for gestational age: Central= 5/31 (20%) Peripheral= 16/73 (30%) p value not reported OR= 0.69 (95% CI 0.23 to 2.07)** Neonatal unit care: Central= 5/31 (20%) Peripheral= 45/73 (83%) p value not reported OR= 0.12 (95% CI 0.04 to 0.35)** Neonatal unit admissions were for hypoglycemia (32%), polycythemia (14%), jaundice (5%), and respiratory distress (5%) 'There was no significant difference in HbA_{1C} acheived in central compared with peripheral hospital sites'	made within the design or analysis to balance the comparison groups for potential confounders - No A3 Groups were comparable at baseline, including all major confounding and prognostic factors - Unclear B1 Comparison groups received the same care apart from the intervention(s) studied - No B2 Participants receiving care were kept 'blind' to treatment allocation - N/A B3 Individuals administering care were kept 'blind' to treatment allocation - N/A C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	BMI > 30 kg/m2 = 18% Booking HbA_{1c}: ≤7% = 51% Mean booking HbA_{1c}= 7.8% (SD 1.8*) Mean booking HbA_{1c} for women with type 1 diabetes= 7.5%* (SD 1.7*) Mean booking HbA_{1c} for women with type 2 diabetes= 7.0%* (SD 2.1*) Mode of delivery: Vaginal and/or operative vaginal= 57%* Emergency caesarean section= 25%* Elective caesarean section= 18%* Ethnicity: Caucasian= 90%* Indo, Asian or African= 10%* In the non-Caucasian group, 1 woman had type 1 diabetes and the other 9 had type 2 diabetes* Prepregnancy care: 28% of women received prepregnancy care 65% of those seen centrally attended a formal			(actual data not reported) **Calculated by the NCC-WCH based on results reported in the paper	in each group? - N/A C2 b. Groups were comparable for treatment completion - Yes C3 a. For how many participants in each group were no outcome data available? - None C3 b. Groups were comparable with respect to the availability of outcome data - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - N/A D5 Investigators were kept 'blind' to other important confounding and prognostic factors - N/A

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	prepregnancy care clinic 14% of those seen peripherally attended a formal prepregnancy care clinic Folic acid uptake= 43% * indicates information or data reported in Dunne (2012) 'ATLANTIC DIP: Pregnancy outcomes for women with type 1 and type 2 diabetes' which reported on the same cohort of women Inclusion criteria Established diabetes for greater than 6 months before the index pregnancy Exclusion criteria None reported				Other information
Full citation Hadden,D.R., How to improve prognosis in type 1 diabetic pregnancy: Old problems, new concepts, Diabetes Care, 22, B104-B108, 1999	Sample size 856 pregnancies (number of women not reported) Characteristics Not reported Inclusion criteria	Interventions Centralised care (n= 386*) Referred into centralised care during pregnancy (n= 80) Peripheral care (n= 390**) * The total number of	Details Three groups of women were compared: 1) Those who received care at a regional centre throughout pregnancy (centralised) 2) Those who were referred from a peripheral hospital to	Results Caesarean section rate 'not greatly different' between the three groups (no data reported). Live births: Centralised= 331/386* (86%)	Limitations There are conflicting data reported in the paper (see 'Results' section). NICE guidelines

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Ref Id 179754 Country/ies where the study was carried out Northern Ireland Study type Retrospective observational study Aim of the study To review the prognosis of pregnancy in women with type 1 diabetes Study dates 1985 to 1995 Source of funding None reported	Type 1 diabetes Exclusion criteria None reported	pregnancies in this group is reported in the paper as 336. However, this conflicts with the sum of the number of live births (n= 331), stillbirths (n= 9), and abortions (n= 46). Therefore, this is assumed to be a typographical error that should read 386. ** The total number of pregnancies in this group is reported in the paper as 391. However, this conflicts with the sum of the number of live births (n= 347), stillbirths (n= 11), and abortions (n= 32). Therefore, this is assumed to be a typographical error that should read 390.	the regional centre during pregnancy (referred) 3) Those who received care at a peripheral hospital throughout pregnancy (peripheral) It is not clear where the data came from or how they were analysed - 'further analysis of the Belfast data' is the only detail given The range and/or mean gestational age at which women were referred to centralised care was not reported. The reasons for referral were not reported.	Referred= 70/80 (88%) Peripheral= 347/390** (89%) p value not reported Still births: Centralised= 9/386* (2%) Referred= 5/80 (6%) Peripheral= 11/390** (3%) p value not reported OR for centralised vs. peripheral= 0.82 (95% CI 0.34 to 2.01)*** Abortions (not specified whether miscarriage is included in this total; reported as 'abortion' in the paper but may include terminations): Centralised= 46/386* (12%) Referred= 5/80 (6%) Peripheral= 32/390** (8%) p value not reported Neonatal deaths: Centralised= 1/386* (<1%) Referred= 1/80 (1%) Peripheral= 5/390** (1%) p value not reported OR for centralised vs. peripheral= 0.20 (95% CI 0.02 to 1.72)*** Perinatal mortality (per 1,000): Centralised= 25.9	manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation to treatment groups was unrelated to potential confounding factors - Unclear A2 Attempts were made within the design or analysis to balance the comparison groups for potential confounders - No A3 Groups were comparable at baseline, including all major confounding and prognostic factors - Unclear B1 Comparison groups received the same care apart from the intervention(s) studied - Unclear B2 Participants receiving care were kept 'blind' to treatment allocation - N/A B3 Individuals administering care were kept 'blind' to treatment allocation

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				Referred= 75.0 Peripheral= 33.5 Whole of Northern Ireland= 9.3 p value not reported Total fetal loss (per 100): Centralised= 14.0 (calculated as 54/386 by the NCC-WCH) Referred= 13.0 (calculated as 10/80 by the NCC-WCH) Peripheral= 12.1 (calculated as 47/390 by the NCC-WCH) p value not reported OR for centralised vs. peripheral= 1.19 (95% CI 0.78 to 1.80)*** * The total number of pregnancies in this group is reported in the paper as 336. However, this conflicts with the sum of the number of live births (n= 331), stillbirths (n= 9), and abortions (n= 46). Therefore, this is assumed to be a typographical error that should read 386. ** The total number of pregnancies in this group is reported in the paper as 391. However, this conflicts with the sum of the number of live births (n= 347), stillbirths	- N/A C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? - N/A C2 b. Groups were comparable for treatment completion - Yes C3 a. For how many participants in each group were no outcome data available? - None C3 b. Groups were comparable with respect to the availability of outcome data - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				(n= 11), and abortions (n= 32). Therefore, this is assumed to be a typographical error that should read 390. ***Calculated by the NCC-WCH based on results reported in the paper	D4 Investigators were kept 'blind' to participants' exposure to the intervention - N/A D5 Investigators were kept 'blind' to other important confounding and prognostic factors - N/A Other information
Full citation Traub,A.I., Harley,J.M., Cooper,T.K., Maguiness,S., Hadden,D.R., Is centralized hospital care necessary for all insulin-dependent pregnant diabetics?, British Journal of Obstetrics and Gynaecology, 94, 957-962, 1987 Ref Id 224491 Country/ies where the study was carried out Northern Ireland Study type Retrospective observational study	Sample size 221 pregnancies in 187 women Characteristics Mean age (years): Centralised= 27.5 Referred= 26.0 Peripheral= 26.7 P value not reported Mean duration of diabetes (years): Centralised= 13.6 Referred= 9.5 Peripheral= 10.2 P value not reported	Interventions Centralised care (60 pregnancies in 56 women) Referred into centralised care during pregnancy (61 pregnancies in 51 women) Peripheral care (100 pregnancies in 80 women)	Details A variety of methods were used to trace and cross-reference names and hospital numbers to ensure all pregnancies were documented. Other sources of data included admission summaries in the labour wards and special care nurseries, personal recollection by obstetricians and clinicians, labour ward records, congenital abnormality records, diabetic clinic and medical outpatient records. Three groups of women were compared: 1) Those who received care at a regional centre throughout pregnancy (centralised)	Results Caesarean section rate: Centralised= 44% (calculated as 26/60 by the NCC-WCH) Referred= 52% (calculated as 32/61 by the NCC-WCH) Peripheral= 61% (calculated as 61/100 by the NCC-WCH) p value not reported OR for centralised vs. peripheral= 0.49 (95% CI 0.26 to 0.94)* Mean gestational age at delivery was 36.6 weeks 'there was no difference between the three groups' - the data were not reported for each of the three groups.	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation to treatment groups was unrelated to potential confounding factors - Unclear A2 Attempts were made within the design or analysis to balance the comparison groups for potential confounders - No A3 Groups were comparable at baseline, including

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Aim of the study To assess the outcomes of all pregnancies in insulin depedent diabetic women in a single area, including an assesssment of the value of centralising care. Study dates 1979 to 1983 Source of funding None reported	Vascular complications: Centralised= 12.5% Referred= 8% Peripheral= 7% P value not reported Previous perinatal mortality: Centralised= 5.0% Referred= 20.0% Peripheral= 12.0% P value not reported Inclusion criteria All known pregnancies in insulin-dependent diabetic women Exclusion criteria Women who were treated with insulin during pregnancy but discontinued it after delivery		2) Those who were referred from a peripheral hospital to the regional centre during pregnancy (referred) 3) Those who received care at a peripheral hospital throughout pregnancy (peripheral) The range and mean gestational age at which women were referred to centralised care was not reported. The reasons for referral were not reported. No statistical analysis of the data was reported.	Livebirth: Centralised= 54/60 (90%) Referred= 50/61 (82%) Peripheral= 88/100 (88%) p value not reported Miscarriage: Centralised= 4/60 (7%) Referred= 3/61 (5%) Peripheral= 10/100 (10%) p value not reported OR for centralised vs. peripheral= 0.64 (95% CI 0.19 to 2.15)* Stillbirth: Centralised= 0/60 (0%) Referred= 6/61 (10%) Peripheral= 2/100 (2%) p value not reported OR for centralised vs. peripheral= 0.33 (95% CI 0.02 to 6.90)* Early neonatal death (out of total number of live births as reported in paper): Centralised= 1/54 (2%) Referred= 0/50 (0%) Peripheral= 1/88 (1%) p value not reported Late neonatal death (out of total number of live births as reported in paper): Centralised= 1/54 (2%)	all major confounding and prognostic factors - Unclear B1 Comparison groups received the same care apart from the intervention(s) studied - Unclear B2 Participants receiving care were kept 'blind' to treatment allocation - N/A B3 Individuals administering care were kept 'blind' to treatment allocation - N/A C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? - N/A C2 b. Groups were comparable for treatment completion - Yes C3 a. For how many participants in each group were

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				Referred= 0/50 (0%) Peripheral= 1/88 (1%) p value not reported Total neonatal deaths (combination of early and late neonatal death, out of all women)*: Centralised= 2/60 (3%) Referred= 0/61 (0%) Peripheral= 2/100 (2%) OR for centralised vs. peripheral= 1.69 (95% CI 0.23 to 12.32)* Infant death (out of total number of live births as reported in paper): Centralised= 0/54 (0%) Referred= 1/50 (2%) Peripheral= 1/88 (1%) p value not reported Perinatal mortality (rate/1000 births) Centralised= 18.5 Referred= 107 Peripheral= 33.3 p value not reported Total fetal loss (including abortions, stillbirths, and deaths within 1 year of life): Centralised= 7.1% (calculated as 4/60 by the NCC-WCH) Referred= 15.5% (calculated	no outcome data available? - None C3 b. Groups were comparable with respect to the availability of outcome data - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - N/A D5 Investigators were kept 'blind' to other important confounding and prognostic factors - N/A Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				as 9/61 by the NCC-WCH) Peripheral= 5.5% (calculated as 6/100 by the NCC-WCH) p value not reported OR for centralised vs. peripheral= 1.12 (95% CI 0.30 to 4.14)* Birthweight > 95th centile occured in 3.3% of pregnancies. The data were not reported for each of the three groups. Mean birthweight was 3368g 'there was no difference between the three groups' - the data were not reported for each of the three groups. *Calculated by the NCC-WCH based on results reported in the paper	

A.6 What are the target ranges for blood glucose in women with type 1, type 2 or gestational diabetes during pregnancy?

Study details	Participants	Methods	Results	Comments
Full citation Rowan,J.A., Gao,W., Hague,W.M., McIntyre,H.D., Glycemia and its relationship to outcomes in the	Population Women aged between 18 and 45 years who developed gestational diabetes mellitis (GDM).	Methods The original trial was a prospective randomised multicentre study. Baseline glycaemia was	Main outcomes Outcomes based on postprandial glucose Pre-eclampsia, n/N Group 1 (< 6.4mmol/l): 19/486 Group 2 (> 6.4mmol/l): 26/238	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential

Study details	Participants	Methods	Results	Comments
metformin in gestational diabetes trial, Diabetes Care, 33, 9-16, 2010 Ref Id 240556 Design Secondary analysis of RCT Country/ies where the study was carried out Australia and New Zealand Aim of study To determine how glucose control influenced trial outcomes in the original MiG trial, to assess the influence of additional baseline factors and to examine differences between treatment arms at different levels of glycaemia. Study dates October 2002 to November 2006 Funding Original trial supported by grants from: The Auckland Medical Research Foundation National Women's Evelyn Bond Charitable Trust Health Research Council of New Zealand National Health and Medical Research Council of Australia	Sample size N = 751 enrolled: 733 had data collected 724 had available glucose data By fasting blood glucose level: ≤ 5.3mmol/l: n = 486 > 5.3mmol/l: n = 240 By blood glucose level: < 6.4mmol/l: n = 486 > 6.4mmol/l: n = 238 Interventions Original trial Intervention: Metformin Control: Insulin Baseline characteristics Age Reported in the original study but not in the context of secondary analysis Body mass index (BMI), n < 25kg/m2 = 131 (18%) 25 to 29kg/m2 = 183 (25%) ≥ 30kg/m2 = 419 (57%) Ethnicity, n European Caucasian/mixed = 373	measured using an oral glucose tolerance test (OGTT) and HbA_{1c} at randomisation to treatment. Alignment with DCCT values for HbA_{1c} was not reported. Treatment glycaemia was measured using capillary glucose readings taken four times daily (fasting and two hours after the start of each meal). Means were calculated separately for each participant. Out of 733 women for whom data were collected: 7 did not have FPG 8 did not have postprandial glucose 9 had no measurements recorded 724 women were included in this secondary analysis. A composite indicator of neonatal morbidity included neonatal hypoglycaemia (≥ 2 glucose readings < 2.6mmol/l), respiratory distress (> 4 hours respiratory support), need for phototherapy, birth trauma, 5 minute Apgar score < 7 or premature birth (< 37 weeks' gestation)	RR = 0.36 (95% CI 0.30 to 0.43)* LGA, n/N Group 1 (< 6.4mmol/l): 56/486 Group 2 (> 6.4mmol/l): 59/238 RR = 0.46 (95% CI: 0.33 to 0.64)* *Calculated by NCC-WCH technical team; dichotomised between second and third tertiles (6.4mmol/l) as the cut-off between tertiles one and two was considered to be very near normal blood glucose levels and therefore too tight for diabetic women. Outcomes based on fasting glucose Pre-eclampsia, n/N Group 1 (≤ 5.3mmol/l): 57/486 Group 2 (> 5.3mmol/l): 59/240 RR = 0.48 (95% CI 0.35 to 0.67)* LGA, n/N Group 1 (≤ 5.3mmol/l): 22/486 Group 2 (> 5.3mmol/l): 23/240 RR = 0.47 (95% CI: 0.27 to 0.83)* *Calculated by NCC-WCH technical team; dichotomised between second and third tertiles (5.3mmol/l).	confounding factors. No - randomisation was not carried out with respect to blood glucose targets. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. Yes - confounders entered into multiple logistic regression models. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. No - diabetes treatment varied as participants were randomised to metformin or insulin in the original trial. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A - secondary analysis. B3: Individuals administering care were kept 'blind' to treatment allocation. N/A - secondary analysis. C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear. C2: a. How many participants did not complete treatment in each group? N/A - secondary

Study details	Participants	Methods	Results	Comments
	(51%) Polynesian = 156 (21%) Asian/other = 204 (28%) Nulliparity, n Yes = 233 (32%) No = 500 (68%) History of pre-eclampsia, n Yes = 55 (7%) No = 445 (61%) Nulliparity = 233 (32%) History of LGA, n Yes = 162 (22%) No = 338 (46%) Nulliparity = 233 (32%) Maternal familial history of diabetes, n Yes = 343 (47%) No = 390 (53%) P-values only reported with respect to outcome. Inclusion criteria Aged between 18 and 45 years Received a diagnosis of GDM according to the Australasian Diabetes in Pregnancy Society (ADIPS)	Large for gestational age (LGA) was defined as > 90th percentile. A definition of pre-eclampsia was not provided. Treatment administered in response to monitoring was not reported. Statistical analyses Mean glucose measures were assessed as continuous variables and categorised quartiles and tertiles. Tertiles were chosen for reporting purposes to give larger group sizes. Bivariable analysis of baseline characteristics was undertaken to explore outcome associations. The Breslow-Day method was used to assess interactions with glycaemic control via stratified analysis and logistic regression. Multivariable logistic regression was used to identify independent risk factors associated with neonatal composite outcome and maternal pre-eclampsia.		analysis. b. The groups were comparable for treatment completion. N/A - secondary analysis. C3: a. For how many participants in each group were no outcome data available? Overall 27 out of 751 (3.6%) enrolled into the original trial (missing data). b. The groups were comparable with respect to the availability of outcome data. Unclear. D. Detection bias D1: The study had an appropriate length of follow-up. Yes. D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome. Yes. D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A - secondary analysis. D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A - secondary analysis. Other information None.

Study details	Participants	Methods	Results	Comments
	Pregnant with a single foetus between 20 and 33 weeks of gestation Met the hospital's usual criteria for starting insulin treatment After lifestyle advice had more than one capillary blood glucose measurement > 5.4mmol/l Exclusion criteria Pre-pregnancy diagnosis of diabetes Contraindication for metformin Foetal anomaly Gestational hypertension Pre-eclampsia Foetal growth restriction Ruptured membranes	Backward stepwise multinomial logistic regression was used to investigate associations between potential risk factors and birth weight, categorised into small for gestational age (SGA), appropriate for gestational age (AGA) and large for gestational age .		
Full citation Landon,M.B., Gabbe,S.G., Piana,R., Mennuti,M.T., Main,E.K., Neonatal morbidity in pregnancy complicated by diabetes mellitus: predictive value of maternal glycemic profiles, American Journal of Obstetrics and Gynecology, 156, 1089-1095, 1987 Ref Id 216952 Design	Population Pregnant diabetic women who delivered at the Hospital of the University of Pennsylvania. Sample size N = 75 By blood glucose level: < 110mg/dl (6.1mmol/l): n = 43 > 110mg/dl: n = 32	Methods Maternal and neonatal charts of 75 diabetic women who delivered between 1982 and 1984 were reviewed. All patients used glucose self-monitoring after initial antepartum evaluation at 12 weeks' gestation. Optimal glucose was considered to be < 100mg/dl (5.5mmol/l) for fasting plasma glucose and < 120mg/dl (6.6mmol/l) for pre-	Main outcomes Mean HbA_{1c} during third trimester, ± SD < 110mg/dl: 5.9 ± 0.9 > 110mg/dl: 7.5 ± 1.1 Mean difference = -1.6* (95% CI -2.1 to -1.1)*†# Mode of delivery < 110mg/dl: Caesarean = 20 (8 primary, 12 repeat) Vaginal = 23 > 110mg/dl:	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. No. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes.

Study details	Participants	Methods	Results	Comments
Retrospective chart review Country/ies where the study was carried out United States of America Aim of study To assess the relationship between glycaemic control and perinatal morbidity in women with type 1 diabetes. Study dates 1982 to 1984 Funding Not reported.	Interventions No specific intervention. Mean capillary blood glucose dichotomised according to level of control achieved: < 110mg/dl considered optimal > 110mg/dl considered sub-optimal Baseline characteristics P-values not reported. Mean age, years ± SD < 110mg/dl: 27 ± 3 > 110mg/dl: 29 ± 5 Mean pre-pregnancy weight, kg ± SD < 110mg/dl: 59.0 ± 10.0 > 110mg/dl: 61.7 ± 10.9 Duration of diabetes, years ± SD < 110mg/dl: 11.3 ± 6 > 110mg/dl: 12.7 ± 8 Pre-eclampsia, n (%) < 110mg/dl: 9 (21.0) > 110mg/dl: 6 (18.7) Inclusion criteria No specific inclusion	prandial blood glucose. Patients obtained glucose measurements at least four times daily. Mean capillary glucose was determined from a minimum of 16 weeks of measurements. A total of 68 patients had readings for the entire second and third trimester. Seven women were admitted to hospital during the second trimester due to low blood glucose. Patients were followed up weekly as outpatients. All infants were initially observed in NICU. Specific treatments administered in response to monitoring were not reported. Glycaemic control was determined by HbA1 (rather than HbA_{1c}) during the third trimester. HbA1 values were therefore not DCCT-aligned. Mode of delivery was either vaginal or Caesarean. Perinatal morbidity included large for gestational age	Caesarean = 16 (7 primary, 9 repeat) Vaginal = 16 RR = 0.93 (95% CI 0.58 to 1.49)* LGA, n/N < 110mg/dl: 4/43 > 110mg/dl: 11/32 RR = 0.27 (95% CI 0.09 to 0.77)* *Calculated by the NCC-WCH technical team †Adjusted using t-distribution due to small sample size #Values were reported as HbA1. Mean HbA_{1c} values were calculated as 5.4% (< 110mg/dl) and 6.8% (> 110mg/dl). It was not possible to convert standard deviations therefore mean differences were calculated using HbA1 values.	B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Unclear. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear. C2: a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion. N/A C3: a. For how many participants in each group were no outcome data available? Overall 7 out of 75 (9.3%) across trimesters two and three. b. The groups were comparable with respect to the availability of outcome data. Unclear. D. Detection bias D1: The study had an appropriate length of follow-up. Yes.

Study details	Participants	Methods	Results	Comments
	criteria were defined. Exclusion criteria Not reported	(LGA) which was defined as birth weight > 90th percentile. Statistical analyses Methods used: For categorical variables X2 contingency tests with Yate's correction or Fisher's exact tests were used as appropriate For continuous variables Student's t-tests were sued Linear regression was used to assess the relationship between mean capillary blood glucose and HbA_{1c}.		D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome. Yes. D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A Other information None.
Full citation Combs,C.A., Gunderson,E., Kitzmiller,J.L., Gavin,L.A., Main,E.K., Relationship of fetal macrosomia to maternal postprandial glucose control during pregnancy., Diabetes Care, 15, 1251-1257, 1992 Ref Id 261442 Design Retrospective review (prospective data) Country/ies where the study was carried out	Population Consecutive pregnant women with pre-existing diabetes enrolled into the Diabetes and Pregnancy Program of the University of California. Sample size N = 111 By blood glucose level: < 7.8mmol/l: n = 66 > 7.8mmol/l: n = 45	Methods 111 consecutive pregnant women admitted to the study were assessed for foetal macrosomia at delivery. Women were White class B to RF. All women were seen weekly or biweekly as outpatients. Patients were instructed to measure blood glucose at least four times daily (one fasting, three post-prandial). In order to reach target values diet plans were	Main outcomes Macrosomia, n/N Postprandial glucose < 7.8mmol: 14/66* Postprandial glucose > 7.8mmol: 18/45* RR = 0.53 (95% CI 0.29 to 0.95)† *Values from weeks 29 to 32 of gestation only based on significance in multiple logistic regression ($\beta = 1.76 \pm 0.82$, $p < 0.05$). †Calculated by the NCC-WCH technical team. Categories of postprandial blood glucose were dichotomised by the NCC-WCH	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. Yes - potential confounders included in multiple regression. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. No.

Study details	Participants	Methods	Results	Comments
United States of America Aim of study To assess factors that contribute to macrosomia in infants of diabetic mothers. Study dates November 1981 to August 1989 Funding Not reported.	Interventions No specific intervention. Women targeted to reach the following blood glucose values: Fasting < 5.9mmol/l (105mg/dl) Postprandial < 7.8mmol/l (140mg/dl) Baseline characteristics Data and p-values were not presented with respect to glucose levels. Mean maternal age, years ± SD Macrosomia: 30.1 ± 4.8 No macrosomia: 31.0 ± 5.2 Mean age at onset of diabetes, years ± SD Macrosomia: 18.4 ± 8.3 No macrosomia: 19.3 ± 9.5 Mean BMI, kg/m2 ± SD Macrosomia: 25.2 ± 5.6 No macrosomia: 26.2 ± 7.2 Nulliparity, n/N (%) Macrosomia: 13/32 (42)	devised for each woman based on energy needs, insulin therapy and nutrients for pregnancy. Treatment administered in response to monitoring was not reported. Women were divided into two groups for analysis: Foetal macrosomia No macrosomia Foetal macrosomia was defined as > 90th percentile for sex and gestational age based on California norms. Three infants were delivered by Caesarean section due to being small for gestational age. Alignment with DCCT values for HbA_{1c} was not reported. Statistical analyses For univariate analyses: X2 for categorical variables Two-tailed Student's t-test for continuous variables P < 0.05 was considered significant. Stepwise multiple logistic regression was used to identify associations between macrosomia and several	technical team according to the target value set for treatment by the study authors of < 7.8mmol/l. This is not exact as a value of 7.84mmol/l was used to separate the relevant categories.	B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear. C2: a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion. N/A C3: a. For how many participants in each group were no outcome data available? N/A - retrospective analysis. b. The groups were comparable with respect to the availability of outcome data. N/A - retrospective analysis. D. Detection bias D1: The study had an appropriate length of follow-up. Yes.

Study details	Participants	Methods	Results	Comments
	No macrosomia: 47/79 (61) Inclusion criteria Diagnosis of diabetes mellitus established before pregnancy Enrollment in the program before 12 weeks' gestation Delivery after 36 weeks' gestation Exclusion criteria Women with gestational diabetes	predictor variable combinations.		D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome. Yes. D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A Other information None.
Full citation Sacks,D.A., Feig,D.S., Liu,I.L., Wolde-Tsadik,G., Managing type I diabetes in pregnancy: how near normal is necessary?, Journal of Perinatology, 26, 458-462, 2006 Ref Id 234259 Design Randomised controlled trial Country/ies where the study was carried out United States of America Aim of study	Population Pregnant women with type 1 diabetes who presented for prenatal care before 13 weeks' gestation. Sample size N = 22 By blood glucose levels: Rigid targets: n = 13 Less rigid targets: n = 9	Methods Eligible women were recruited into the study. Identification of type 1 diabetes was made based on insulin requirements or history of abrupt onset of diabetes, DKA or both. All participants were instructed in diet, insulin administration and glucose self-monitoring. Women were to record blood glucose seven times per day, before and after each meal	Main outcomes Mean HbA_{1c}, % ± SD 1st trimester Rigid targets: 6.3 ± 0.7 Less rigid targets: 7.5 ± 1.5 Mean difference = -1.2 (95% CI -2.32 to -0.08)* 2nd trimester Rigid targets: 5.6 ± 0.8 Less rigid targets: 6.1 ± 0.6 Mean difference = -0.5 (95% CI -1.12 to 0.12)* 3rd trimester Rigid targets: 5.9 ± 0.6	Limitations NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups. Yes. A2: There was adequate concealment of allocation. N/A. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes - though small groups therefore analyses likely underpowered.

Study details	Participants	Methods	Results	Comments
To determine patient compliance and to report preliminary findings. Study dates April 2000 to March 2003 Funding Not reported. Laboratory analyses donated by Quest Diagnostics. Glucose meters, software and technical support donated by Roche Diagnostics.	Interventions Rigid targets: Fasting values 60 to 90mg/dl (3.3 to 5.0mmol/l) Postprandial values 120 to 140mg/dl (6.7 to 7.8mmol/l) Less rigid targets: Fasting values 95 to 115mg/dl (5.3 to 6.4mmol/l) Postprandial values 155 to 175mg/dl (8.6 to 9.7mmol/l) Baseline characteristics Mean age, years ± SD Rigid targets: 32.5 ± 5.5 Less rigid targets: 31.2 ± 3.9 P-value = 0.86 Mean pre-pregnancy BMI, kg/m2 ± SD Rigid targets: 24.0 ± 2.8 Less rigid targets: 28.7 ± 5.9 P-value = 0.05 Ethnicity, % caucasian Rigid targets: 77 Less rigid targets: 67 P-value = 0.66 Nulliparity, %	and at bedtime. Allocation was carried out using computer-generated block randomisation. HbA_{1c} measurements were repeated once each trimester. Alignment with DCCT values for HbA_{1c} was not reported. During the intrapartum period maternal blood glucose was maintained between 70 to 110mg/dl. Treatment administered in response to monitoring was not reported. Outcomes were as follows: Mean maternal HbA_{1c} Mode of delivery (vaginal or Caesarean) Mean birth weight Statistical analyses Sample size calculation Hypothesised treatment difference between groups (rate of hypoglycaemia) was 19% minus 5% = 14%. Level of significance = 0.05 Power = 80%	Less rigid targets: 6.2 ± 0.8 Mean difference = -0.3 (95% CI -0.95 to 0.35)* Mode of delivery, n/N Rigid targets: Caesarean = 8/13 (5 elective, 3 emergency) Less rigid targets: Caesarean = 6/9 (2 elective, 6 emergency) RR = 1.08 (95% CI 0.57 to 2.04)* *Calculated by the NCC-WCH technical team.	B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear. C2: a. How many participants did not complete treatment in each group? 4 out of 13 in the less rigid group. b. The groups were comparable for treatment completion. No. C3: a. For how many participants in each group were no outcome data available? 4 out of 13 in the less rigid group. b. The groups were comparable with respect to the availability of outcome data. No. D. Detection bias D1: The study had an appropriate length of follow-up. Yes.

Study details	Participants	Methods	Results	Comments
	Rigid targets: 62 Less rigid targets: 56 P-value = 1.00 Inclusion criteria Type 1 diabetes Presented to prenatal care before 13 weeks' gestation Exclusion criteria Not reported.	Implied sample size of 84 patients per group. Analytical methods Fisher's exact test for categorical data Non-parametric Wilcoxon's ranksum test for continuous data P-values < 0.05 were deemed significant.		D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome. Yes. D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A Other information Reasons for the loss of four patients from the less rigid group: Two had first trimester spontaneous abortions One deleted because participated in the study with an earlier pregnancy One declined to attend appointments
Full citation Demarini,S., Mimouni,F., Tsang,R.C., Khoury,J., Hertzberg,V., Impact of metabolic control of diabetes during pregnancy on neonatal hypocalcemia: a	Population Pregnant women with type 1 diabetes (White classification B to RT) and their infants.	Methods Eligible women were randomly assigned to either treatment group. All women received twice daily insulin injections and dietary regulation and measured their blood glucose at least	Main outcomes Mean HbA_{1c} in the first trimester, % ± SD Strict control: 9.4 ± 1.9† Customary control: 9.4 ± 1.8† MD = 0.0 (95% CI -0.62 to 0.62)*#	Limitations NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups. Unclear - randomisation

Study details	Participants	Methods	Results	Comments
randomized study, Obstetrics and Gynecology, 83, 918-922, 1994 Ref Id 261563 Design Randomised controlled trial Country/ies where the study was carried out United States of America Aim of study To test the hypothesis that strict glycaemic control during pregnancy reduces the risk of neonatal hypocalcaemia in infants of diabetic mothers. Study dates July 1978 to June 1989 Funding Part funded by grants from the National Institutes of Health.	Sample size N = 137 Strict control n = 68 Customary control n = 69 Interventions Intervention Strict management to achieve fasting blood glucose values < 80mg/dl (4.44mmol/l) and 1.5 hour post-prandial blood glucose < 120mg/dl (6.66mmol/l). Control Standard care as practised in the community to achieve fasting blood glucose values < 100mg/dl (5.55mmol/l) and post-prandial blood glucose < 140mg/dl (7.77mmol/l). Baseline characteristics Mean maternal age, years ± SD Strict control: 25.3 ± 5.0 Customary control: 26.6 ± 4.8	twice daily. Women in the strict control group were admitted to hospital immediately at entry into the study in order to achieve blood glucose control. Women in the customary care group were only admitted if targets were not achieved after one week as an outpatient. Women receiving strict glycaemic control were seen weekly. Women in the customary care group were seen bi-weekly in the first and second trimesters and weekly thereafter. In addition to self-monitoring, blood glucose was assessed weekly using glucose reflectance meters. Every four weeks both laboratory and self-monitoring instruments were verified against laboratory instruments. HbA_{1c} was determined using column chromatography. The normal range was based on assay reference values in children. Alignment with DCCT values for HbA_{1c} was not reported.	Mean HbA_{1c} in the second trimester, % ± SD Strict control: 7.8 ± 1.4† Customary control: 7.7 ± 1.4† MD = 0.1 (95% CI -0.37 to 0.57)*# Mean HbA_{1c} in the third trimester, % ± SD Strict control: 7.5 ± 1.2† Customary control: 7.6 ± 1.1† MD = -0.1 (95% CI -0.49 to 0.29)*# *Calculated by the NCC-WCH technical team. #Values were reported as HbA1. It was not possible to convert standard deviations therefore mean differences were calculated using HbA1 values. †Corresponding HbA_{1c} values are as follows: 9.4 = 8.5% 7.5 = 6.8% 7.6 = 6.9% 7.7 = 7.0% 7.8 = 7.1%	methods are not described. A2: There was adequate concealment of allocation. Unclear. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes, though exact p-values were not reported. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. No - women were assessed more frequently in trimesters one and two and were admitted to hospital immediately to achieve glycaemic control. B2: Participants receiving care were kept 'blind' to treatment allocation. Unclear. B3: Individuals administering care were kept 'blind' to treatment allocation. Unclear. C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear. C2: a. How many participants did not complete treatment in each group? Not reported. b. The groups were comparable for treatment completion. Unclear.

Study details	Participants	Methods	Results	Comments
	P-value = not significant Mean parity ± SD Strict control: 0.72 ± 0.92 Customary control: 0.97 ± 0.97 P-value = not significant Mean duration of diabetes, years ± SD Strict control: 11.9 ± 6.1 Customary control: 11.3 ± 7.1 P-value = not significant Exact p-values were not reported unless results were statistically significant. Inclusion criteria A diagnosis of type 1 diabetes. Exclusion criteria Not reported.	Treatment administered in response to monitoring was not reported. Statistical analyses Continuous data were analysed using Student's t-tests and ANOVA. Categorical data were analysed using either Fisher's exact tests or X2 tests.		C3: a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data. Unclear. D. Detection bias D1: The study had an appropriate length of follow-up. Yes. D2: The study used a precise definition of outcome. N/A - mean HbA_{1c} values were reported. D3: A valid and reliable method was used to determine the outcome. Yes, though frequency of testing was not reported. D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear. D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear. Other information None.
Full citation Farrag,O.A., Prospective study of 3 metabolic regimens in pregnant diabetics, Australian and New Zealand Journal of	Population Saudi women with overt insulin dependent diabetes (White class B and C).	Methods Sixty Saudi pregnant women with White class diabetes B or C were recruited to the study during the first trimester of pregnancy.	Main outcomes Maternal hypoglycaemia, n/N < 5.6 SI = 7/16 5.6 to 6.7 SI = 0/29 6.7 to 8.9 SI = 0/15	Limitations A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally

Study details	Participants	Methods	Results	Comments
Obstetrics and Gynaecology, 27, 6-9, 1987 Ref Id 181071 Design Randomised controlled trial Country/ies where the study was carried out Saudi Arabia Aim of the study To determine the best regimen of metabolic control in pregnant women in Saudi Arabia. Study dates Not reported Funding Not reported.	Sample size N = 60 Interventions Women were targeted to achieve the following fasting blood glucose values depending upon the regimen to which they were assigned. Group A (n = 16) < 5.6 SI Group B (n = 29) 5.6 to 6.7 SI Group C (n = 15) 6.7 to 8.9 SI Baseline characteristics Maternal age, range (years) 24 to 40 Parity, range Between 3 and 8 previous children No other baseline characteristics were reported.	All women were admitted to hospital to regualte insulin and dietary requirements. All women received a diet suitable to meet maternal and fetal needs, comprising carbohydrates, protein and fat. Diets consisted of 3 meals and 2 snacks per day with equal carbohydrate distribution. Fasting and post-prandial blood glucose measurements were taken on the third day of the diet. Women were then allocated to one of three treatment regimen aimed at achieving blood glucose of < 5.6 SI (mmol/l), 5.6 to 6.7 SI or 6.7 to 8.9 SI. Randomisation methods were not described. Insulin administration was managed based on one unit per 0.6 SI increase above the targeted value. Blood glucose was checked two days later and insulin therapy adjusted where necessary. Insulin was given as a mixture of NPH and regular insulin half an hour before breakfast (2:1 ratio) and half an hour before dinner (1:1 ratio). Average duration of stay in	RR = 39.71 (95% CI: 2.26 to 697.01)* Pre-eclampsia, n/N < 5.6 SI = 1/16 5.6 to 6.7 SI = 0/29 6.7 to 8.9 SI = 3/15 RR = 0.92 (95% CI: 0.10 to 8.59)* Caesarean section, n/N < 5.6 SI = 2/16 5.6 to 6.7 SI = 3/29 6.7 to 8.9 SI = 6/15 RR = 0.62 (95% CI: 0.15 to 2.64)* Large for gestational age, n/N < 5.6 SI = 0/16 5.6 to 6.7 SI = 0/29 6.7 to 8.9 SI = 13/15 RR = 0.10 (95% CI: 0.006 to 1.68)* Perinatal mortality, n/N < 5.6 SI = 0/16 5.6 to 6.7 SI = 0/29 6.7 to 8.9 SI = 2/15 RR = 0.53 (95% CI: 0.03 to 11.14)* *Calculated by the NCC-WCH technical team. Data were dichotomised between groups A and B (< 5.6 versus ≥ 5.6 SI).	across groups). Unclear A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear - insufficient baseline characteristics were reported. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? None b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Yes

Study details	Participants	Methods	Results	Comments
	Inclusion criteria Women with White class diabetes B or C (insulin dependent) Exclusion criteria Presence of any medical complications other than diabetes Women who presented after the first trimester	hospital was eight days. At 20 and 28 weeks' gestation women were admitted to hospital for re-adjustment of insulin therapy. Large for gestational age was defined as births greater than the 90th percentile. Statistical analyses Not described.		C3: a. For how many participants in each group were no outcome data available? Not reported b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. No - pre-eclampsia, maternal hypoglycaemia and perinatal mortality were not defined. D3: A valid and reliable method was used to determine the outcome. Unclear D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information Both fasting and 2 hour postprandial blood glucose were measured. It is unclear from the methods which of these values targets

Study details	Participants	Methods	Results	Comments
				given to women relate to. It was assumed that targets related to fasting blood glucose due to the low values assigned. The numbers of women who achieved the assigned targets were not reported however mean blood glucose values in each group were as follows: < 5.6 SI = 5.0 SI 5.6 to 6.7 SI = 6.1 SI 6.7 to 8.9 SI = 8.4 SI Four Caesarean sections were elective and seven emergency. Of the elective Caesareans two were for pre-eclampsia, one for a clinically large baby and one for low biochemical results. Of the emergency Caesareans five were due to failure to progress during labour and two due to fetal distress.

A.7 What is the target value for HbA_{1c} in women with type 1, type 2 or gestational diabetes during pregnancy?

Study details	Participants	Methods	Results	Comments
Full citation Barnes,R.A., Edghill,N., Mackenzie,J., Holters,G., Ross,G.P., Jalaludin,B.B.,	Population Women diagnosed with gestational diabetes in a high-risk,	Methods Data from a computerised database were analysed for eligible women. Pre-pregnancy BMI, weight gain,	Main outcomes Large for gestational age OR for HbA_{1c} > 5.5%	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual

Study details	Participants	Methods	Results	Comments
Flack,J.R., Predictors of large and small for gestational age birthweight in offspring of women with gestational diabetes mellitus, Diabetic Medicine, 30, 1040-1046, 2013 Ref Id 305869 Study design Retrospective audit Country/ies where the study was carried out Australia Aim of the study To identify independent predictors of small and large for gestational age infants in women with gestational diabetes mellitus. Study dates August 1992 to April 2009. Source of funding None.	ethnically diverse population of women in Australia. Sample size N = 1695 Interventions No specific intervention. Characteristics Mean gestational age at diagnosis, weeks 28.1 ± 5.3 Mean duration of treatment for GDM, weeks 11.0 ± 5.3 Ethnicity, n (%) South East Asian = 626 (36.7%) Middle Eastern = 467 (27.6%) European = 380 (22.4%) Indian and Pakistani = 146 (8.6%) Samoan = 33 (1.9%) Non-white African = 25 (1.5%) Maori = 18 (1.1%) Inclusion criteria	HbA_{1c} at presentation and treatment modality (diet or insulin) were recorded. Diagnosis of GDM was based on ADIPS criteria using a 75g OGTT: Fasting ≥ 5.5mmol/l 1 hour postprandial ≥ 10.0mmol/l 2 hour postprandial ≥ 8.0mmol/l Therapy comprised diet and insulin was added if the following targets were not met: Fasting glucose < 5.5mmol/l 2 hour postprandial < 7.0mmol/l HbA_{1c} was determined at diagnosis of GDM. Based on the findings of previous studies HbA_{1c} was dichotomised at 5.5% which represented the upper limit of normal in the third trimester. LGA was defined as > 90th percentile adjusted for age, maternal height and weight, parity and ethnicity. Statistical analyses Data were expressed as mean ± SD. Logistic regression was used to identify significant predictors of SGA and LGA infants. Backward selection was used to determine final models. P-values < 0.05 were taken to be statistically significant.	versus ≤ 5.5% = 1.38 (95% CI 1.01 to 1.90)* *Result taken from logistic regression.	A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. Yes A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Unclear B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion (that is, there

Study details	Participants	Methods	Results	Comments
	Singleton pregnancies Diagnosed with GDM by ADIPS criteria Exclusion criteria Incomplete data (except HbA_{1c}) Delivery < 36 weeks' gestation Last clinic weight recorded > 4 weeks before delivery			were no important or systematic differences between groups in terms of those who did not complete treatment). N/A C3: a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). N/A D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A Other information None

Study details	Participants	Methods	Results	Comments
Full citation Ekbom,P., Damm,P., Feldt-Rasmussen,B., Feldt-Rasmussen,U., Jensen,D.M., Mathiesen,E.R., Elevated third-trimester haemoglobin A $_{1c}$ predicts preterm delivery in type 1 diabetes, Journal of Diabetes and its Complications, 22, 297-302, 2008 Ref Id 210981 Aim of study To assess the predictive value of HbA_{1c} for preterm delivery in women with type 1 diabetes. Study design Prospective cohort Country/ies where the study was carried out Denmark Study dates Not reported Funding Not reported.	Population Caucasian women with type 1 diabetes and a living foetus admitted to the study clinic before 17 weeks' gestation. Sample size N = 213 By tertile of HbA_{1c} at 28 weeks' gestation < 6.0%: n = 71 6.0 to 6.5%: n = 60 > 6.5%: n = 82 Interventions No specific intervention. Baseline characteristics Data and p-values were not presented according to HbA_{1c} levels. Mean age, years ± SD Delivery at term: 30 ± 5 Preterm: 29 ± 4 Mean BMI, kg/m2 ± SD	Methods Women entered the study consecutively. Women were asked to perform home blood glucose measurements ≥ 4 times per day. Measurements of HbA_{1c} were performed ≥ 5 times throughout pregnancy. Labour was routinely induced after 38 to 40 weeks of completed gestation. Treatment administered in response to monitoring was not reported. HbA_{1c} values were chosen to represent metabolic control at different time points in pregnancy: 10 weeks = early pregnancy 20 weeks = second trimester 28 weeks = late pregnancy Outcomes were as follows with some definitions given in a previous paper (reference provided by authors): Preterm delivery (< 37 weeks' gestation) Pre-eclampsia (not defined) Large for gestational age (LGA) (not defined) Neonatal hypoglycaemia (not defined) Perinatal mortality (after 22 weeks' gestation or within one week of delivery) HbA_{1c} values DCCT-aligned in 10% of women.	Main outcomes Maternal hypoglycaemic episodes (not defined) by HbA_{1c} measured at 28 weeks' gestation, n/N ≤ 6.5: 22/131 > 6.5: 11/82 RR = 1.08 (95% CI 0.55 to 2.10)* *Calculated by NCC-WCH technical team using a threshold of 6.5%, based on a normal range of 4.1% to 6.4% for non-pregnant individuals quoted in the study. Only one outcome is reported here as all other outcomes of interest were reported in relation to gestational age at delivery not HbA_{1c} values.	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. No. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes - no specific intervention, all participants treated per study centre protocol. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear.

Study details	Participants	Methods	Results	Comments
	Delivery at term: 24 ± 3 Preterm: 24 ± 3 Mean duration of diabetes, years ± SD Delivery at term: 12 ± 8 Preterm: 12 ± 8 Nulliparity, n (%) Delivery at term: 80 (56) Preterm: 35 (49) Inclusion criteria Type 1 diabetes Living foetus Admitted before 17 weeks' gestation Exclusion criteria Microalbuminuria at the first clinic visit Overt nephropathy at the first clinic visit Miscarriages (≤ 22 weeks' gestation) Twin pregnancies	Statistical analyses Two-tailed Student's t-tests were used for continuous variables. Fisher's exact tests and Yate's-corrected X2 were used for categorical data. In repeated comparisons Bonferroni adjustment was made to the nominal p-value. Multivariate logistic regression was used to identify variables independently associated with pre-term delivery.		C2: a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion. N/A C3: a. For how many participants in each group were no outcome data available? None. b. The groups were comparable with respect to the availability of outcome data. Yes. D. Detection bias D1: The study had an appropriate length of follow-up. Yes. D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome. Yes. D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A Other information

Study details	Participants	Methods	Results	Comments
				None.
Full citation Mikkelsen,M.R., Nielsen,S.B., Stage,E., Mathiesen,E.R., Damm,P., High maternal HbA_{1c} is associated with overweight in neonates, Danish Medical Bulletin, 58, A4309-, 2011 Ref Id 247990 Study design Retrospective cohort Country/ies where the study was carried out Denmark Aim of study To determine the prevalence of pregnant women with gestational diabetes who do not obtain optimal HbA_{1c} values before delivery and to assess whether elevated HbA_{1c} increase the risk of LGA. Study dates 2007 Funding Not reported.	Population All women who delivered at the study clinic during the study period who were diagnosed with gestational diabetes. Sample size N = 148 By HbA_{1c} level: Obtained treatment goal (HbA_{1c} ≤ 5.6%): n = 97 Did not obtain (HbA_{1c} > 5.6%): n = 51 Interventions No specific intervention. Treatment goals: ≤ 5.6% considered optimal > 5.6% considered poor Baseline characteristics Mean age, years ± SD ≤ 5.6%: 33.3 ± 4.5	Methods After diagnosis of gestational diabetes women who met inclusion criteria were enrolled in the study. Demographic and clinical details were obtained from original medical records. All women received individualised dietary advice for one hour and were trained in self-monitoring of blood glucose (SMBG). Telephone contact was offered to help achieve goals for SMBG. Treatment goals: SMBG between 4 and 6mmol/l preprandially SMBG 4 and 8mmol/l postprandially HbA_{1c} ≤ 5.6% 97/148 (66%) women obtained the target of a last measured HbA_{1c} ≤ 5.6%. Alignment with DCCT values for HbA_{1c} was not reported. Treatment consisted of a calorie-restricted diet and exercise. Insulin was administered if women had ≥ 2 blood glucose values above the treatment goal within 14 days of commencing treatment.	Main outcomes LGA ≤ 5.6%: 18/97 > 5.6%: 20/51 Study reports adjusted OR using ≤ 5.6% as the referent: OR = 3.12 (95% CI 1.28 to 7.61) Using > 5.6% as the referent: Crude RR = 0.47 (95% CI 0.27 to 0.81)* Pre-eclampsia ≤ 5.6%: 7/97 > 5.6%: 3/48 RR = 1.23 (95% CI 0.33 to 4.56)* Shoulder dystocia ≤ 5.6%: 2/97† > 5.6%: 0/51† RR = 2.65 (95% CI 0.13 to 54.18)* Neonatal hypoglycaemia ≤ 5.6%: 4/97 > 5.6%: 7/51 Study reports adjusted OR using ≤ 5.6% as the referent:	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. Yes - adjusted for confounders in multiple regression. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. No - differed in BMI, OGTT result and HbA_{1c}. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was

Study details	Participants	Methods	Results	Comments
	> 5.6%: 31.2 ± 4.9 P-value = 0.01 Mean pre-pregnancy BMI, kg/m2 ± SD ≤ 5.6%: 27.8 ± 6.5 > 5.6%: 30.9 ± 6.0 P-value = 0.006 Parity > 1, n (%) ≤ 5.6%: 64 (66.0) > 5.6%: 34 (66.7) P-value = 0.81 Ethnicity, n (%) Caucasian: 85 (57.4) Middle East: 37 (25.0) Asia: 11 (7.4) Other: 15 (10.1) P-value = 0.08 Inclusion criteria Diagnosis of gestational diabetes before 34 weeks (OGTT ≥ 9.0mmol/l or FPG > 6.1mmol/l) Singleton pregnancies HbA_{1c} outside the normal range at diagnosis and measured again < 3 weeks before delivery ≥ 3 weeks between	Treatment administered in response to monitoring was not reported. Outcomes were as follows: Frequency of large for gestational age (LGA) infants (birth weight > 90th percentile, adjusted for sex and gestational age) Pre-eclampsia (blood pressure ≥ 140/90mmHg accompanied by proteinuria) Shoulder dystocia (shoulder delivery required obstetrical manoeuvres and downward traction) Neonatal hypoglycaemia (symptomatic or asymptomatic glucose 2 hours postpartum < 2.5mmol/l) Mode of delivery (vaginal and Caesarean) Statistical analyses Continuous data were analysed using Mann-Whitney U tests or Student's t-tests. Binary outcomes were analysed using X2 tests and odds ratios were calculated. Multiple logistic regression was used to investigate potential confounders including: Ethnicity Parity Smoking status	OR = 6.17 (95% CI 1.31 to 29.04) Using > 5.6% as the referent: Crude RR = 0.30 (95% CI 0.15 to 0.60)* Mode of delivery (Caesarean/N) ≤ 5.6%: 32/97 (14 elective, 18 emergency) > 5.6%: 16/51 (5 elective, 11 emergency) RR = 1.05 (95% CI 0.64 to 1.72)* Induction of labour was performed in 50 women who achieved HbA_{1c} ≤5.6% and 33 women who did not achieve this HbA_{1c}. *Calculated by NCC-WCH technical team. †A value of 0.5 was added to each cell in the contingency table in order for a relative risk to be calculated.	adjusted to allow for differences in length of follow-up). Unclear - likely enrolled at different times in gestation. C2: a. How many participants did not complete treatment in each group? None. b. The groups were comparable for treatment completion. Yes. C3: a. For how many participants in each group were no outcome data available? Unclear. b. The groups were comparable with respect to the availability of outcome data. Unclear. D. Detection bias D1: The study had an appropriate length of follow-up. Yes. D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome. Yes. D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A

Study details	Participants	Methods	Results	Comments
	HbA_{1c} measurements Exclusion criteria Missing HbA_{1c} values Malignant disorder	Maternal family history of diabetes Weight gain during pregnancy Pre-pregnancy BMI Maternal age Two-sided p-values of < 0.05 were considered statistically significant.		Other information None.
Full citation Vaarasmaki,Marja S., Hartikainen,Anna Liisa, Anttila,Marjatta, Pramila,Sirkka, Koivisto,Maila, Factors predicting peri- and neonatal outcome in diabetic pregnancy, Early Human Development, 59, 61-70, 2000 Ref Id 280037 Study design Retrospective cohort Country/ies where the study was carried out Finland Aim of study To assess factors associated with adverse perinatal outcomes in pregnant women with type 1 diabetes. Study dates 1986 to 1995 Funding Not reported.	Population Consecutive births to women with type 1 diabetes in a geographically defined catchment area in Finland. Sample size N = 296 By HbA_{1c} level: Optimal glycaemic control: n= 48 Poor glycaemic control: n = 36 Interventions No specific intervention. Baseline characteristics Nulliparity, n/N (%) Optimal control: 18/48 (37.5)	Methods Women in the cohort were from the two northernmost provinces in Finland. Data were obtained from one tertiary hospital and four central secondary hospitals. Data were recorded prospectively. Prior to 1992 optimal HbA_{1c} control was considered to be < 8.0% (based on HbA1 rather than HbA_{1c}), after 1992 optimal control was between 4.0 and 6.0% for HbA_{1c}. An HbA1 of 8.0% corresponds to an HbA_{1c} of 7.3%. Medical history, the course of pregnancy and delivery and neonatal clinical information were recorded. Data from diabetic women were compared to unpublished data on 44 678 singleton pregnancies in non-diabetic women obtained between 1991 and 1995 in the same geographical area. Women were followed up at least every fourth week until 22 weeks' gestation, then at 1 to 2 week intervals until week 36. Thereafter	Main outcomes Neonatal unit stay > 10 days† Optimal control: 2/48 Poor control: 11/36 RR = 0.14 (95% CI 0.03 to 0.59)* †Only one outcome is reported here due to the poor quality of the study; no other studies included in this review assessed neonatal unit stay. *Calculated by NCC-WCH technical team	Limitations NICE checklist for cohort studies, taken from Appendix D of the NICE guidelines manual A. Selection bias A1: The method of allocation to treatment groups was unrelated to potential confounding factors. Unclear. A2: Attempts were made within the design or analysis to balance the comparison groups for potential confounders. No. A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Unclear - unlikely as data obtained from five separate hospitals. B2: Participants receiving care were kept 'blind' to treatment allocation. N/A

Study details	Participants	Methods	Results	Comments
	Poor control: 19/36 (52.8) No p-value reported No data were reported for maternal age, BMI or ethnicity in relation to glycaemic control. Inclusion criteria Singleton pregnancies Gestational age ≥ 22 weeks, or Birth weight ≥ 500g Exclusion criteria Not reported.	visits were twice weekly or women were hospitalised until delivery. Treatment administered in response to monitoring was not reported. All neonates were examined by a paediatrician immediately after delivery. Infants were admitted to a neonatal unit only as a result of medical indications. Outcomes were as follows: Large for gestational age (LGA) (birth weight ≥ 2 SD above the normal mean for gestational age) Perinatal mortality (stillbirths and neonatal deaths before 7 days of life) Observation in the neonatal unit Neonatal hypoglycaemia (blood glucose ≤ 1.7mmol/l more than twice or one low value alongside IV glucose during the first 48 hours of life) Statistical analyses Risk ratios and 95% CIs were calculated for factors recorded at baseline in association with adverse events. Odds ratios for large for gestational age infants were calculated using logistic regression and adjusted for maternal BMI and HbA_{1c} at different stages of pregnancy.		B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear. C2: a. How many participants did not complete treatment in each group? N/A b. The groups were comparable for treatment completion. N/A C3: a. For how many participants in each group were no outcome data available? Unclear - only 84 of 296 pregnancies had glycaemic control data reported, 48 for optimal control and 36 for poor control. b. The groups were comparable with respect to the availability of outcome data. Unclear - see point C3a. D. Detection bias D1: The study had an appropriate length of follow-up. Yes. D2: The study used a precise definition of outcome. Yes. D3: A valid and reliable method was used to determine the outcome. Yes.

Study details	Participants	Methods	Results	Comments
				D4: Investigators were kept 'blind' to participants' exposure to the intervention. N/A D5: Investigators were kept 'blind' to other important confounding and prognostic factors. N/A Other information None.

A.8 What is the effectiveness of blood glucose monitoring in predicting adverse outcomes in women with type 1, type 2, or gestational diabetes during pregnancy?

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation Varner,M.W., Efficacy of home glucose monitoring in diabetic pregnancy, American Journal of Medicine, 75, 592-596, 1983 Ref Id 179004 Country/ies where the study was carried out United States of America Study type Randomised controlled trial Aim of the study To report the maternal and neonatal outcomes comparing glucose monitoring with a conventional outpatient management protocol.	Sample size N = 30 Characteristics Mean maternal age, years Daily: 24.0 ± 4.0 Weekly: 23.3 ± 4.4 Average parity Daily: 1.4 ± 0.7 Weekly: 1.3 ± 0.8 P-values were not reported. Inclusion criteria Required insulin before conception	Interventions Home blood glucose monitoring (n = 15) Weekly blood glucose monitoring (n = 15)	Details A total of 34 women were recruited to the study and 30 agreed to participate. Women were assigned to either the control or experimental group using a random number sequence. Women in the control group were managed according to protocols at the study institution. All women were admitted after the first clinic visit for metabolic control. Glucose targets were fasting of 70 to 110mg/dl and two-hour postprandial of 80 to	Results Caesarean section Daily: 7/14 Weekly: 9/14 RR = 0.78 (95% CI 0.39 to 1.54)* Vaginal delivery Daily: 7/14 Weekly: 5/14 RR = 1.40 (95% CI 0.56 to 3.50)* Neonatal hypoglycaemia Daily: 4/14 Weekly: 7/14 RR = 0.57 (95% CI 0.20 to 1.59)*	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Yes A2: There was

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Study dates February 1980 to 1981. Source of funding Research Fellowship from the Iowa Affiliate of the American Diabetes Association.	At < 20 weeks' gestation Exclusion criteria Not reported.		130mg/dl. Women received a three meal/three snack American Diabetes Association diet based on 35kcal/kg ideal body weight. Subcutaneous insulin was administered twice daily as regular plus NPH or lente intermediate-acting insulin. Once metabolic control was established women were discharged and followed in the high-risk obstetric unit. Women were then seen every two weeks until 32 weeks' gestation then weekly thereafter. Serum glucose (fasting, two hours after breakfast and two hours after lunch) were measured on one day each week. Insulin was adjusted accordingly. Women were instructed to telephone on a weekly basis to report glucose levels and any complications. All women were admitted for the remainder of the pregnancy after 36 weeks' gestation. Women in the experimental group were also admitted for metabolic control after the first clinic visit. During admission women were instructed in the use of a home-monitoring system for whole blood glucose determination. Women	*Calculated by the NCC-WCH technical team.	adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			were discharged when metabolic control had been established and followed in the high-risk obstetric unit. Women were also instructed to telephone on a weekly basis. Fasting plus two-hour postprandial morning, afternoon and evening blood glucose values were monitored daily by the women. One woman from each group had a spontaneous first trimester miscarriage therefore were excluded from analyses. Outcomes included mode of delivery, weeks' gestation and weight at birth. Perinatal morbidity was assessed by polycythaemia, hypocalcaemia, hyperbilrubinaemia and hypoglycaemia. Neonatal hypoglycaemia was defined as serum glucose < 30mg/dl. Statistical analyses were performed using either small-sample t-tests or the X2 test. P-values < 0.05 were taken to be statistically significant.		equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes C2: a. How many participants did not complete treatment in each group? One b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Yes C3: a. For how many participants in each group were no outcome data available? One b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					whom outcome data were not available). Yes D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. Yes - though assisted vaginal delivery was not reported separately. D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information None

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation Bancroft,K., Tuffnell,D.J., Mason,G.C., Rogerson,L.J., Mansfield,M., A randomised controlled pilot study of the management of gestational impaired glucose tolerance, BJOG: An International Journal of Obstetrics and Gynaecology, 107, 959-963, 2000 Ref Id 257978 Country/ies where the study was carried out UK Study type Randomised controlled trial Aim of the study To undertake a pilot study for a trial to determine whether less intensive management of impaired glucose intolerance in pregnancy is beneficial Study dates Not reported Source of funding None reported	Sample size 68 women Characteristics Age at delivery (years) Self-monitoring= 29.7 +/- 6.23 No self-monitoring= 31.9 +/- 5.17 p value not reported BMI at booking (kg/m2) Self-monitoring= 31.2 +/- 6.7 No self-monitoring= 27.5 +/- 6.1 p value not reported Ethnicity: Asian Self-monitoring= 10/32 (31%) No self-monitoring= 11/36 (31%) p value not reported Caucasian Self monitoring= 22/32 (69%) No self-monitoring= 25/36 (69%) p value not reported Family history of type 2 diabetes Self-monitoring= 12/32 (37%) No self-monitoring= 11/36 (31%) p value not reported	Interventions Self-monitoring (n= 32) No self-monitoring (n= 36)	Details Ethics committee approval was obtained (it was not reported from whom). Written informed consent was obtained from all participants. Women were recruited from two specialist diabetic/antenatal clinics after referral from general antenatal clinics. Glucose tolerance tests were performed at the discretion of individual clinicians. Women were randomly assigned to one of two groups by a computer generated code, stratified by trimester of diagnosis and ethnicity. Randomisation was administered by telephone from a trial centre. The diabetologist was aware of the woman's group allocation, but the obstetrician was kept blind. All women were given dietary advice about restricting carbohydrate intake to 185 grams per day. In one group, women had their glucose metabolism monitored by means of capillary glucose series (1 to 2 hours after meals) 5 times a week, with glycosylated	Results Vaginal birth Self-monitoring= 22/32 (69%) No self-monitoring= 25/36 (69%) p value not significant Caesarean section Self-monitoring= 10/32 (31%) No self-monitoring= 11/36 (31%) p value not significant HbA_{1c} (%): 28 weeks (n= 8 in each group) Self-monitoring= 4.9 +/- 0.7 No self-monitoring= 5.5 +/- 1.1 p value not significant 32 weeks (n= 20 in monitored group, n= 19 in unmonitored group) Self-monitoring= 5.2 +/- 0.8 No self-monitoring= 5.0 +/- 1.3 p= 0.03 36 weeks (n= 31 in monitored group, n= 32 in unmonitored group) Self-monitoring= 5.3 +/- 0.8 No self-monitoring= 5.6	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials A1 An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups) - Yes A2 There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation) - Yes A3 The groups were comparable at baseline, including all major confounding and prognostic factors – Not clear B1 The comparison groups received the same care apart from the intervention(s) studied - Yes B2 Participants receiving care were

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Gestation at entry to study (weeks) Self-monitoring= 31 (range 24 to 38) No self-monitoring= 32 (range 15 to 37) p value not reported HbA_{1c} at entry to study (%) Self-monitoring= 5.3 +/- 0.83 No self-monitoring= 5.6 +/- 0.96 p value not reported Fasting glucose (mmol/L) Self-monitoring= 4.6 (range 3.5 to 5.8) No self-monitoring= 4.7 (range 3.5 to 7.0) p value not reported 2 hour glucose (mmol/L) Self-monitoring= 8.5 (range 7.9 to 10.8) No self-monitoring= 8.9 (range 7.8 to 11.0) p value= 0.025 Inclusion criteria Women with impaired glucose tolerance (fasting blood glucose level <7.0 mmol/L and 2 hour blood glucose between 7.8 mmol/L and 11 mmol/L)		haemoglobin measurements performed monthly (self-monitoring group). Insulin was started if 5 or more capillary glucose measurements were > 7.0 mmol/L in one week. Women in the other group did not have their glucose metabolism monitored, although they also had monthly glycosylated haemoglobin measurements (no self-monitoring group). Groups were compared using Student's t test or Mann-Whitney U test, Fisher's exact test or Pearson $\chi 2$ test. A p value of < 0.05 was used to indicate significance.	+/- 1.2 p value not significant 38 weeks (n= 24 in monitored group, n= 27 in unmonitored group) Self-monitoring= 5.3 +/- 0.9 No self-monitoring= 5.5 +/- 0.9 p value not significant At term (n= 10 in each group) Self-monitoring= 5.1 +/- 0.8 No self-monitoring= 5.5 +/- 0.9 p value not significant Birthweight > 90th centile for gestation Self-monitoring= 8/32 (25%) No self-monitoring= 7/36 (19%) p value not significant Neonatal hypoglycaemia Self-monitoring= 2/32 (6%) No self-monitoring= 6/36 (17%) p value not significant Shoulder dystocia Self-monitoring= 0/32 No self-monitoring= 1/36 p value not reported	kept 'blind' to treatment allocation - No B3 Individuals administering care were kept 'blind' to treatment allocation - Yes C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? - None C2 b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment) - Yes C3 a. For how many participants in each group were no outcome data available? - None C3 b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Exclusion criteria			Stillbirths Self-monitoring= 0/32 No self-monitoring= 0/36 Neonatal deaths Self-monitoring= 0/32 No self-monitoring= 0/36	systematic differences between groups in terms of those for whom outcome data were not available). - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - Unclear D5 Investigators were kept 'blind' to other important confounding and prognostic factors - No Other information
Full citation Espersen,T., Klebe,J.G., Self-monitoring of blood glucose in pregnant diabetics. A comparative study of the blood glucose level and course of pregnancy in pregnant diabetics on an out-patient regime before and after the introduction of methods for	Sample size 121 women Characteristics White classification: B Self-monitoring= 17 No self-monitoring= 19 C	Interventions Self-monitoring (n= 61) No self-monitoring (n= 62)	Details Two types of self-monitoring systems were used - a reflectometer (Aimes) with Dextrostix test strips, and Haemoglucotest 1-44 test strips. The distribution of the two systems was based on the limited number of each	Results Large for gestational age (>90th percentile) Self-monitoring= 12/61 (20%) No self-monitoring= 19/62 (31%) p value not significant	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation to treatment groups was unrelated to potential

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
home analysis of blood glucose, Acta Obstetricia et Gynecologica Scandinavica, 64, 11-14, 1985 Ref Id 234547 Country/ies where the study was carried out Denmark Study type Cohort study Aim of the study To determine whether self-monitoring of blood glucose is better than no self-monitoring Study dates 1978 to 1981 Source of funding None reported	Self-monitoring= 12 No self-monitoring= 17 D Self-monitoring= 23 No self-monitoring= 21 FR Self-monitoring= 23 No self-monitoring= 21 All pregnancies were singleton pregnancies No other characteristics were reported Inclusion criteria Women with type 1 diabetes Exclusion criteria Women in White group A		type of equipment. Women in one group were taught to self-monitor blood glucose (self-monitoring group). They were asked to test 5 times a day (7am, 10am, 1pm, 4pm, and 8pm) at least twice a week. Women were seen at an out-patients' clinic once every 1 or 2 weeks. Adjustments were made to the amount of insulin given, if necessary. The other group was made up of women who did not use self-monitoring (no self-monitoring group).		confounding factors – Yes A2 Attempts were made within the design or analysis to balance the comparison groups for potential confounders – No A3 Groups were comparable at baseline, including all major confounding and prognostic factors – Unclear B1 Comparison groups received the same care apart from the intervention(s) studied – Yes B2 Participants receiving care were kept 'blind' to treatment allocation – No B3 Individuals administering care were kept 'blind' to treatment allocation – Unclear C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					each group? - None C2 b. Groups were comparable for treatment completion - Yes C3 a. For how many participants in each group were no outcome data available? - None C3 b. Groups were comparable with respect to the availability of outcome data - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention – No D5 Investigators were kept 'blind' to other important confounding and prognostic factors - No Other information
Full citation	Sample size	Interventions	Details	Results	Limitations

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Goldberg,J.D., Franklin,B., Lasser,D., Jornsay,D.L., Hausknecht,R.U., Ginsberg-Fellner,F., Berkowitz,R.L., Gestational diabetes: impact of home glucose monitoring on neonatal birth weight, American Journal of Obstetrics and Gynecology, 154, 546-550, 1986 Ref Id 218186 Country/ies where the study was carried out USA Study type Retrospective case control study Aim of the study To determine the effect of home glucose monitoring compared to weekly clinic monitoring on neonatal outcomes Study dates July 1979 to July 1984 Source of funding Supported in part by National Institutes of Health Grant HD 11583 and the Sosnoff Foundation.	116 women Characteristics Age (years) Daily monitoring= 30.4 +/- 6 Weekly monitoring= 30.1 +/- 6 p value not significant Ethnicity: Hispanic Daily monitoring= 64% Weekly monitoring= 59% p value not significant Black Daily monitoring= 33% Weekly monitoring= 34% p value not significant Gestational age at time of diagnosis Daily monitoring= 26.8 +/- 7 weeks Weekly monitoring= 29.1 +/- 7 weeks p value not significant Oral glucose tolerance test: Fasting (mg/dL) Daily monitoring= 98 +/- 17 Weekly monitoring= 104 +/- 16 p < 0.05 1 hour (mg/dL) Daily monitoring= 206 +/- 41	Daily monitoring (n= 58) Weekly monitoring (n= 58)	Before 1983, all pregnant women were screened for glucose intolerance with a 3 hour oral glucose tolerance test if they had 1 of 10 risk factors (previously published but not stated in paper). After 1983, all women were given a 50g oral glucose screening test. Women with an oral glucose plasma value of ≥135 mg/dL after 1 hour had a full 100g oral glucose tolerance test. The diagnosis of glucose intolerance was based on the criteria for O'Sullivan and Mahan modified to correct for the methodologic change from the Somogyi-Nelson method to glucose oxidase and for measurements of plasma rather than whole blood glucose. The diagnosis of gestational diabetes was made when two values met or exceeded: fasting 95mg/dL, 1 hour 180mg/dL, 2 hour 155mgt/dL, or 3 hour 135 mg/dL. All women were started on a diabetic diet (30 to 35 kilocalories per kilogram of ideal body weight; 25% fat,	Vaginal birth Daily monitoring= 27/58 (47%) Weekly monitoring= 37/58 (65%) p value not significant Forceps Daily monitoring= 12/58 (21%) Weekly monitoring= 5/58 (10%) p value not significant Caesarean section Daily monitoring= 18/58 (32%) Weekly monitoring= 14/58 (25%) p value not significant Large for gestational age (not defined) Daily monitoring= 7 (12%) Weekly monitoring= 24 (41%) p<0.005 Compliance with daily glucose monitoring was >90%	NICE guidelines manual. Appendix E: Methodology checklist: Case-control studies 1.1 The study addresses an appropriate and clearly focused question – Well covered 1.2 The cases and controls are taken from comparable populations - Adequately covered 1.3 The same exclusion criteria are used for both cases and controls - Well covered 1.4 What was the participation rate for each group (cases and controls)? – Not applicable 1.5 Participants and non-participants are compared to establish their similarities or differences – Not reported 1.6 Cases are clearly defined and differentiated from controls - Well covered 1.7 It is clearly established that controls are not cases

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Weekly monitoring= 200 +/- 37 p value not significant 2 hour (mg/dL) Daily monitoring= 182 +/- 43 Weekly monitoring= 177 +/- 45 p value not significant 3 hour (mg/dL) Daily monitoring= 138 +/- 44 Weekly monitoring= 127 +/- 42 p value not significant Inclusion criteria Not reported Exclusion criteria Women registering after 36 weeks		25% protein, and 50% complex carbohydrate). All women were seen weekly in the clinic, where a 2 hour postprandial capillary blood glucose measurement was performed. Before September 1983, women did not undertake home glucose monitoring (Weekly monitoring group). After September 1983, all women that were enrolled were started on home glucose monitoring (Daily monitoring group). Fasting and 1 hour postprandial values were obtained daily using a visually read Chemstrip bG glucose test (Bio-Dynamics, Indianapolis, Indiana). These women were randomly listed by computer (method not described) and matched with the women in the study group for age, prepregnancy weight, height, ideal body weight, and parity (primiparas or multiparas). Insulin therapy was begun if fasting glucose values were >95 mg/dL (at home) or if postprandial values were >120mg/dL (at home or in the clinic). Statistical analysis was performed with use of a		- Well covered 1.8 Measures were taken to prevent knowledge of primary exposure from influencing case ascertainment – Not reported 1.9 Exposure status is measured in a standard, valid, and reliable way - Well covered 1.10 The main potential confounders are identified and taken into account in the design and analysis – Adequately covered 1.11 Have confidence intervals been provided? – Not applicable Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			two-tailed t test and McNemar's test to assess significance.		
Full citation Hawkins,J.S., Casey,B.M., Lo,J.Y., Moss,K., McIntire,D.D., Leveno,K.J., Weekly compared with daily blood glucose monitoring in women with diet-treated gestational diabetes, Obstetrics and Gynecology, 113, 1307-1312, 2009 Ref Id 240657 Country/ies where the study was carried out USA Study type Retrospective cohort study Aim of the study To determine whether daily monitoring reduces macrosomia compared to weekly office testing in women with gestational diabetes Study dates January 1991 to March 2001 Source of funding None reported	Sample size 990 women Characteristics Age (years) Daily monitoring= 29.9 +/- 5.8 Weekly monitoring= 29.4 +/- 5.6 p=0.15 Ethnicity: White Daily monitoring= 9/315 (2.9%) Weekly monitoring= 43/675 (6.4%) African American Daily monitoring= 24/315 (7.6%) Weekly monitoring= 75/675 (11.1%) Hispanic Daily monitoring= 272/315 (86.3%) Weekly monitoring= 531/675 (78.7%) Other Daily monitoring= 10/315 (3.2%) Weekly monitoring= 26/675 (3.9%) p value for ethnicity overall= 0.023	Interventions Daily monitoring (n= 315) Weekly monitoring (n= 675)	Details The study was deemed exempt from ethical review by the Institutional Review Board of the University of Texas Southwestern Medical Center. Women were screened for gestational diabetes between 24 and 28 weeks of gestation. They were given a 50g oral glucose screening test (Allegiance Healthcare Corp., McGaw Park, IL). If their serum glucose exceeded 140 mg/dL (but was less than 200 mg/dL) at 1 hour, they were given a 100g 3 hour oral glucose tolerance test after an overnight fast. Women with two or more abnormal values according to the National Diabetes Data Group thresholds were diagnosed with gestational diabetes. Women whose 50g glucose screening test exceeded 200mg/dL underwent a fasting capillary blood glucose measurement. If their glucose value was less than 105 mg/dL then they underwent a 100g glucose	Results Vaginal delivery (including forceps delivery) Daily monitoring= 199/315 (63.2%) Weekly monitoring= 453/675 (67.1%) p= 0.22 Forceps delivery Daily monitoring=7/315 (2.2%) Weekly monitoring= 25/675 (3.7%) p= 0.22 Caesarean section Daily monitoring= 116/315 (36.8%) Weekly monitoring= 222/675 (32.9%) p= 0.22 Shoulder dystocia Daily monitoring= 5/315 (1.6%) Weekly monitoring= 13/675 (1.9%) p= 0.71 Large for gestational age Daily monitoring= 73/315 (23.1%) Weekly monitoring=	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation to treatment groups was unrelated to potential confounding factors – Yes A2 Attempts were made within the design or analysis to balance the comparison groups for potential confounders – Yes A3 Groups were comparable at baseline, including all major confounding and prognostic factors – No B1 Comparison groups received the same care apart from the intervention(s) studied – Unclear B2 Participants receiving care were kept 'blind' to treatment allocation – No

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Gestational age at diagnosis of diabetes (weeks) Daily monitoring= 25.3 +/- 6.2 Weekly monitoring= 26.5 +/- 5.6 p= 0.003 50g glucose challenge test (mg/dL) Daily monitoring= 179 +/- 41 Weekly monitoring= 171 +/- 37 p= 0.005 100g glucose tolerance test (md/dL): Fasting blood sugar Daily monitoring= 99 +/- 20 Weekly monitoring= 99 +/- 16 p=0.86 1 hour Daily monitoring= 210 +/- 31 Weekly monitoring= 209 +/- 33 p=0.60 2 hour Daily monitoring= 186 +/- 36 Weekly monitoring= 188 +/- 36 p= 0.45 3 hour		tolerance test. All women were managed in a special morning obstetrics clinic held weekly at a hospital. They received dietary counselling, including instructions to limit daily caloric intake to 35 kilocalories per kilogram of body weight and which foods to avoid. All women underwent monitoring of serum fasting glucose during each weekly office visit. From January 1998, women were given a self-monitoring blood glucose meter (Accucheck Advantage or Advantage IIm Boehringer Mannheim Corp, Indianapolis, IN) upon diagnosis of gestational diabetes. These women were instructed to test their capillary blood glucose four times a day (preprandially, including a morning fasting value and before bedtime) (Daily group). The pregnancy outcomes of these women were compared to the women who were diagnosed with gestational diabetes prior to January 1998, who did not receive a blood glucose meter and relied whose serum fasting	232/675 (34.4%) p < 0.001 (This difference remained significant after adjustment for maternal demographic variables and gestational age at diagnosis) Neonatal hypoglycaemia Daily monitoring= 23/315 (7.3%) Weekly monitoring= 30/675 (4.4%) p= 0.06 Women with home glucose monitors (daily monitoring group) measured their glucose an average of 3.7 +/- 0.7 times a day.	B3 Individuals administering care were kept 'blind' to treatment allocation – No C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? - None C2 b. Groups were comparable for treatment completion - Yes C3 a. For how many participants in each group were no outcome data available? - None C3 b. Groups were comparable with respect to the availability of outcome data - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Daily monitoring= 139 +/- 36 Weekly monitoring= 143 +/- 37 p= 0.18 Inclusion criteria Women with diet treated gestational diabetes and who had risk factors for gestational diabetes (included family history of diabetes, personal history of gestational diabetes, prior delivery of a stillborn, malformed or macrosomic neonate) Singleton pregnancies Exclusion criteria Noncephalic gestations Women with persistent fasting glucose of 105 or greater		glucose was measured at weekly office visits (Weekly group). Large for gestational age ≥ 90th percentile birth weight for gestational age distribution (population specific) Statistical analysies performed include $\chi 2$, Student t test, and multiple logistic regressions. Values of $p < 0.05$ were considered statistically significant.		D4 Investigators were kept 'blind' to participants' exposure to the intervention - No D5 Investigators were kept 'blind' to other important confounding and prognostic factors - Unclear Other information For women who were pregnant between January 1991 and December 1996, only those with risk factors for gestational diabetes (including family history of diabetes, personal history of gestational diabetes, prior delivery of a stillborn, malformed or macrosomic neonate) were screened for gestational diabetes. From January 1997 all pregnant women were routinely screened for gestational diabetes between 24 and 28 weeks, however, only women who also had risk factors for gestational diabetes were included in this study to minimise selection bias.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation Manderson,J.G., Patterson,C.C., Hadden,D.R., Traub,A.I., Ennis,C., McCance,D.R., Preprandial versus postprandial blood glucose monitoring in type 1 diabetic pregnancy: a randomized controlled clinical trial, American Journal of Obstetrics and Gynecology, 189, 507-512, 2003 Ref Id 234197 Country/ies where the study was carried out UK Study type Randomised controlled trial Aim of the study To compare preprandial and postprandial capillary glucose monitoring in pregnant women with type 1 diabetes Study dates Not reported Source of funding Supported by grants from the Department of Health and Social Services, Northern Ireland, the Northern Ireland Mother and Baby Appeal, the Metabolic Unit Research Fund, Royal Victoria Hospital, Belfast,	Sample size 61 women Characteristics Age Preprandial monitoring= 29.7 +/- 4.9 years Postprandial monitoring= 30.0 +/- 4.9 years p= 0.80 BMI (kg/m2) Preprandial monitoring= 25.9 +/- 3.9 Postprandial monitoring= 28.6 +/- 5.8 p= 0.04 Onset of diabetes Preprandial monitoring= 16.4 +/- 9.2 years Postprandial monitoring= 18.0 +/- 10.1 years p= 0.53 All participants had diabetes before pregnancy Initial glycosylated haemoglobin (%) Preprandial monitoring= 7.6 +/- 1.1 Postprandial monitoring= 7.4 +/- 1.4 p= 0.63 All participants were white	Interventions Preprandial monitoring (n= 31) Postprandial monitoring (n= 30)	Details The study was ethically approved (it is not stated who gave ethical approval). Written consent was obtained from the women. At 16 weeks of gestation, women were randomly assigned to one of two monitoring protcols (method of randomisation not reported). Allocations were via a sealed enveloped system, which women selected from a box at the clinic visit. There was a limit of 40 women in each group. Women used a single memory-based glucose reflectance meter (One Touch profile, Lifescan, Inc, Milpitas, Calif). One group of women was asked to monitor before breakfast and preprandially (preprandial monitoring group) and the other group was asked to monitor before breakfast and 1 hour after the commencement of each meal (postprandial monitoring group). During any hospitalisation, women were monitored according to their group assignment. Women were transferred to	Results Caesarean section Preprandial monitoring= 21/31 (68%) Postprandial monitoring= 14/30 (47%) p= 0.10 Neonatal hypoglycaemia (glucose < 1.7 mmol/L during first 72 hours of life or requirement of intravenous glucose treatment) Preprandial monitoring= 9/31 (29%) Postprandial monitoring= 8/30 (26.7%) p value not significant Glycosylated haemoglobin (%): Initial Preprandial monitoring= 7.6 +/- 1.1 Postprandial monitoring= 7.4 +/- 1.4 p= 0.63 Final Preprandial monitoring= 6.3 +/- 0.7 Postprandial monitoring= 6.0 +/- 0.8 p= 0.11 Change from booking Preprandial monitoring= -1.3 +/- 1.0 Postprandial monitoring=	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials A1 An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups) – Not clear A2 There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation) - Yes A3 The groups were comparable at baseline, including all major confounding and prognostic factors - No B1 The comparison groups received the same care apart from the intervention(s) studied - Yes B2 Participants

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
the Royal Maternity Hospital, Royal Victorial Hospital, Belfast, and the Irish Perinatal Society.	Inclusion criteria Women attending or referred to the Regional Joint Metabolic/Antenatal Clinic before 14 weeks' gestation Exclusion criteria Women with a history of hypertension, proteinuric renal disease before pregnancy, or who had a urinary albumin greater than 20g/dL or an albumin/creatinine ratio greater than 2.0mg/mmol at < 20 weeks' gestation were excluded		a four-times daily basal bolus insulin regimen, if not already on this. Insulin doses were adjusted to achieve targets suggested by the American Diabetes Association. Neonatal hypoglycaemia was defined as a blood glucose less than 1.7 mmol/L (analysed at 1 hour after delivery via heel prick). Groups were compared using independent samples t tests (after logarithmic transformation for nonnormally distributed variables), and $\chi 2$ analysis with Yates' correction or Fisher exact test where appropriate. All tests were conducted at the 5% level of significance.	-1.4 +/- 1.3 p= 0.59 Stillbirth Preprandial monitoring= 1/32* Postprandial monitoring= 0/30 p value not reported *This woman was excluded from other analyses Birthweight > 90 percentile Preprandial monitoring= 18/31 (58%) Postprandial monitoring= 15/30 (50%) p= 0.71 Length of stay in neonatal unit (days) Preprandial monitoring= 6.0 (2 to 8) Postprandial monitoring= 4.0 (2 to 12) p= 0.86 Compliance with the monitoring schedule did not differ significantly between the two groups	receiving care were kept 'blind' to treatment allocation - No B3 Individuals administering care were kept 'blind' to treatment allocation - No C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? – 13 women were excluded from the analysis (see 'other information' below), but it is not clear from which group C2 b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment) – Not clear C3 a. For how many participants in each group were no outcome data

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					available? – Not clear C3 b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). – Not clear D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention – Not clear D5 Investigators were kept 'blind' to other important confounding and prognostic factors – Not clear Other information Adherence to the monitoring schedule was low - 47.6% and 30.2% in the preprandial group in

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					trimester 2 and trimester 3 respectively, and 39.7% and 35.7% in the postprandial group in trimester 2 and trimester 3 respectively. There was no significant difference in adherence between the two groups. 13 women were excluded from the analysis - 1 woman withdrew from the study, 3 women had incomplete results, 4 women had spontaneous abortions, 1 woman had a stillbirth, 4 women delivered infants with major congenital abnormalities (leaving 61 women in the analysis)
Full citation de Veciana,M., Major,C.A., Morgan,M.A., Asrat,T., Toohey,J.S., Lien,J.M., Evans,A.T., Postprandial versus preprandial blood glucose monitoring in women with gestational diabetes mellitus requiring insulin therapy, New England Journal	Sample size 66 women Characteristics Age Preprandial monitoring= 31 +/- 6 Postprandial monitoring= 29 +/- 5	Interventions Preprandial monitoring (n= 33) Postprandial monitoring (n= 33)	Details The study was approved for the institutional review boards of the University of California at Irvine and Long Beach Memorial Medical Center. Women with risk factors for gestational diabetes (including body weight >	Results Caesarean section Preprandial monitoring= 13/33 (39%) Postprandial monitoring= 8/33 (24%) RR 1.6 (95% CI 0.8 to 3.4) p= 0.29	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials A1 An appropriate method of randomisation was used to allocate

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
of MedicineN.Engl.J.Med., 333, 1237-1241, 1995 Ref Id 257662 Country/ies where the study was carried out USA Study type Randomised controlled trial Aim of the study To determine whether postprandial or preprandial monitoring is more effectiv in achieving glycaemic control in women with gestational diabetes Study dates Not reported Source of funding None reported	p value not significant Ethnicity: Hispanic Preprandial monitoring= 27/33 (82%) Postprandial monitoring= 29/33 (88%) p value not significant White Preprandial monitoring= 4/33 (12%) Postprandial monitoring= 3/33 (9%) p value not significant Black or Asian Preprandial monitoring= 2/33 (6%) Postprandial monitoring= 1/33 (3%) p value not significant Plasma glucose (mg/dL): At 1 hour Preprandial monitoring= 216 +/- 56 Postprandial monitoring= 214 +/- 67 p value not significant Fasting (at time of 3 hour oral glucose tolerance test) Preprandial monitoring= 137 +/- 38 Postprandial monitoring= 145 +/- 50 p value not significant		120 percent of ideal value, age ≥ 35 years, glucosuria on dipstick urinalysis [≥2+], a history of diabetes in first degree relatives, and a previous unexplained stillbirth or miscarriage) were screened at their initial visits. If the initial screening was normal, these women were also screened at 24 to 28 weeks of gestation. Women without risk factors for gestational diabetes were screened at 24 to 28 weeks. Initial screening involved a measurement of plasma glucose one hour after 50g oral glucose. If the plasma glucose test result was between 140mg/dL and 190 mg/dL, a 3 hour oral glucose tolerance test was done. Gestational diabetes was diagnosed if women had any two of the following plasma glucose values: fasting >105 mg/dL, 1 hour >190 mg/dL, 2 hours >165 mg/dL, 3 hours >145mg/dL. All women with elevated fasting values at the time of the 3 hour test were immediately started on insulin therapy. All other women were initially treated with diet and monitored with weekly fasting and postprandial	Large for gestational age Preprandial monitoring= 14/33 (42%) Postprandial monitoring= 4/33 (12%) RR 3.5 (95% CI 1.3 to 9.5) p= 0.01 Shoulder dystocia Preprandial monitoring= 6/33 (18%) Postprandial monitoring= 1/33 (3%) RR 6.0 (95% CI 0.8 to 47.1) p= 0.10 Neonatal hypoglycaemia Preprandial monitoring= 7/33 (21%) Postprandial monitoring= 1/33 (3%) RR 7.0 (95% CI 0.9 to 53.8) p= 0.05 Stillbirth Preprandial monitoring= 1/33 (3%) Postprandial monitoring= 0/33 (0%) RR not reported p= 1.00	participants to treatment groups (which would have balanced any confounding factors equally across groups) - Yes A2 There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation) – Not clear A3 The groups were comparable at baseline, including all major confounding and prognostic factors - Yes B1 The comparison groups received the same care apart from the intervention(s) studied - Yes B2 Participants receiving care were kept 'blind' to treatment allocation - No B3 Individuals administering care were kept 'blind' to treatment allocation - Unclear C1 All groups were followed up for an

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Week of gestation at diagnosis Preprandial monitoring= 22.9 +/- 7.5 Postprandial monitoring= 21.8 +/- 6.5 p value not significant Inclusion criteria Women with gestational diabetes requiring insulin at or before 30 weeks of gestation Singleton pregnancies Exclusion criteria Women with a history of diabetes before pregnancy Women with pre-existing hypertension, renal disease, or autoimmune disorders		measurements of plasma glucose. If the plasma glucose test result was 190 mg/dL or higher, a 3 hour glucose tolerance test was not performed. Insulin therapy was started in any woman (regardless of 3 hour glucose tolerance test result) if values exceeded 105 mg/dL fasting or 140 mg/dL postprandial. Women were assigned to a group for the duration of their pregnancies using permuted-block randomisation. One group required daily monitoring of fasting, preprandial, and bedtime capillary-blood glucose concentrations (Preprandial group). The other group required daily monitoring of blood glucose concentrations before breakfast (fasting), and one hour after each meal (Postprandial group). If women were hospitalised during pregnancy, women were monitored according to their group assignment. Women measured their blood glucose concentrations using memory-based reflectance glucometers, with all values recorded. Both groups were	A review of patient records of home monitoring during the last four weeks of pregnancy showed similar levels of compliance (≥95%) and achievement of target blood glucose values in the two groups (although women in the postprandial group received more insulin that the women in the preprandial group).	equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? - None C2 b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment) - Yes C3 a. For how many participants in each group were no outcome data available? - None C3 b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			prescribed a diet with a daily allocation of 30 to 35 kilocalories per kilogram of ideal body weight. 40 to 45% of energy was provided by carbohydrates. Calorie intake and food choices were adjusted at weekly visits if needed. Women receiving insulin therapy had their dose adjusted to aim to achieve a fasting blood glucose value of 60 to 90mg/dL and preprandial values of 60 to 105 mg/dL or postprandial values below 140mg/dL. Hypoglycaemia was defined as blood glucose concentration ≤ 30 mg per deciliter Shoulder dystocia was defined when one or more manoeuvres were needed to facilitate vaginal delivery of the neonate's shoulders Infants were assigned birth-weight percentiles according to gestational age and sex with use of the population-specific standards published in California Mann-Whitney U test was used for normally distributed data. Two-tailed Fisher's exact test was used for categorical data.		precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - Unclear D5 Investigators were kept 'blind' to other important confounding and prognostic factors - Unclear Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			Relative risks and 95% confidence intervals were calculated with Epi Info software (version 5, Stone Mountain, Ga).		
Full citation Weisz,B., Shrim,A., Homko,C.J., Schiff,E., Epstein,G.S., Sivan,E., One hour versus two hours postprandial glucose measurement in gestational diabetes: a prospective study, Journal of Perinatology, 25, 241-244, 2005 Ref Id 257977 Country/ies where the study was carried out Israel Study type Prospective observational study Aim of the study To compare outcomes in women with gestational diabetes monitored by 1 hour postprandial glucose measurements to those monitored by 2 hour postprandial glucose measurements Study dates May 1999 to April 2000	Sample size 112 women Characteristics Age (years) 1 hour postprandial monitoring= 30.9 +/- 5.44 2 hour postprandial monitoring= 33.1 +/- 5.24 p= 0.03 Glucose challenge test (50g) 1 hour postprandial monitoring= 169.1 +/- 34.6 2 hour postprandial monitoring= 171.0 +/- 26.7 p value not significant Oral glucose tolerance test (100g): At time of test 1 hour postprandial monitoring= 90.4 +/- 12.0 2 hour postprandial monitoring= 94.8 +/- 13.8 p value not significant At 60 minutes 1 hour postprandial monitoring= 205.3 +/- 27.8 2 hour postprandial monitoring= 210.3 +/- 21.9 p value not significant	Interventions 1 hour postprandial monitoring (n= 66 women) 2 hour postprandial monitoring (n= 46 women)	Details The study was approved by the Sheba Medical Center Institutional Review Board. Women were diagnosed with gestational diabetes based on the Carpenter and Coustan criteria. Women were referred to a diabetes in pregnancy program from two different outpatient clinics in the same city, although both clinics were staffed by the same team of health care professionals. Women seen in one treatment centre were managed by 1 hour postprandial measurements (1 hour postprandial monitoring group), whilst women in the other centre were managed by 2 hour postprandial measurements (2 hour postprandial monitoring group). All women were seen by a registered dietitian for individualised counselling. Women were placed on 1800-2200 calories a day - 40 to 45% carbohydrates,	Results Caesarean section 1 hour postprandial monitoring= 15/66 (24%) 2 hour postprandial monitoring= 14/46 (30%) p= 0.62 Large for gestational age (not defined) 1 hour postprandial monitoring= 5/66 (7.4%) 2 hour postprandial monitoring= 7/46 (15.2%) p value not significant	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation to treatment groups was unrelated to potential confounding factors – Yes A2 Attempts were made within the design or analysis to balance the comparison groups for potential confounders – Unclear A3 Groups were comparable at baseline, including all major confounding and prognostic factors – No B1 Comparison groups received the same care apart from the intervention(s) studied – Yes B2 Participants receiving care were kept 'blind' to treatment allocation –

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Source of funding Supported by a grant from the General Clinical Research Center branch of the National Center for Research Resources (2M01-RR-349)	At 120 minutes 1 hour postprandial monitoring= 174.0 +/- 24.3 2 hour postprandial monitoring= 178.8 +/- 29.5 p value not significant At 180 minutes 1 hour postprandial monitoring= 109.9 +/- 37.2 2 hour postprandial monitoring= 116.9 +/- 40.2 p value not significant Inclusion criteria Not reported Exclusion criteria Women with pregestational diabetes Women with fasting glucose levels of 105 mg/dL or above Twin pregnancies		20% protein, and ≤40% fat. All women were given a memory-based blood glucose meter (One Touch Profile, LifeScan, Inc.) and were asked to measure capillary blood glucose. Glucose levels were measured at fasting and either 1 hour (target value of <140mg/dL) or 2 hours (target value of <120 mg/dL) postprandially. Insulin therapy was initiated if fasting levels exceeded 95 mg/dL (both groups) or target values in more than 30% of measurements. Statistical analysis was performed using Student's t-test, $\chi 2$, and multiple regressions. Stastical significance was set at $p<0.05$.		No B3 Individuals administering care were kept 'blind' to treatment allocation – No C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? – 6 women were lost to follow up, but it is not clear from which group C2 b. Groups were comparable for treatment completion - Unclear C3 a. For how many participants in each group were no outcome data available? - Unclear C3 b. Groups were comparable with respect to the availability of outcome data - Unclear D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - Unclear D5 Investigators were kept 'blind' to other important confounding and prognostic factors - Unclear Other information 6 women were lost to follow up
Full citation Langer,O., Rodriguez,D.A., Xenakis,E.M., McFarland,M.B., Berkus,M.D., Arrendondo,F., Intensified versus conventional management of gestational diabetes, American Journal of Obstetrics and Gynecology, 170, 1036-1046, 1994 Ref Id 236280 Country/ies where the study was carried out USA Study type Prospective cohort study Aim of the study	Sample size 2461 women Characteristics Age (years) 4 times daily monitoring= 30.4 +/- 6 7 times daily monitoring= 30.2 +/- 4 p value not significant Ethnicity: Black 4 times daily monitoring= 3.0% 7 times daily monitoring= 4.1% p value not significant	Interventions 4 times daily monitoring group (n= 1316) 7 times daily monitoring group (n= 1145)	Details All pregnant women were screened for carbohydrate intolerance at 24 to 28 weeks of gestation using a 1 hour glucose challenge. It plasma glucose was ≥ 130 mg/dL, a 3 hour 100g oral glucose tolerance test was done. Gestational diabetes was diagnosed by means of the National Diabetes Data Group glucose threshold. Test results in which one or more values were elevated were considered abnormal. Pregnant women were assigned to clinics in	Results Caesarean section 4 times daily monitoring= 283/1316 (21.5%) 7 times daily monitoring= 172/1145 (15.0%) p value reported as significant (actual value not reported) Large for gestational age 4 times daily monitoring= 265/1316 (20.1%) 7 times daily monitoring= 150/1145 (13.1%) p<0.0001	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation to treatment groups was unrelated to potential confounding factors – Yes A2 Attempts were made within the design or analysis to balance the comparison groups for potential confounders – Unclear

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
To determine whether intensified management of gestational diabetes reduces adverse outcomes Study dates July 1989 to April 1993 Source of funding None reported	White 4 times daily monitoring= 15.0% 7 times daily monitoring= 15.5% p value not significant Hispanic 4 times daily monitoring= 81.0% 7 times daily monitoring= 79.0% p value not significant Other 4 times daily monitoring= 1.0% 7 times daily monitoring= 1.4% p value not significant Obesity (defined as > 27.3 kg/m2) 4 times daily monitoring= 50.0% 7 times daily monitoring= 48.0% p value not significant Previous gestational diabetes 4 times daily monitoring= 16.4% 7 times daily monitoring= 15.2% p value not significant Family history of diabetes 4 times daily monitoring= 45.3%		random order (method of randomisation not reported). Women were assigned to groups on the basis of the availability of memory-based reflectance meters - after a woman already enrolled in the study gave birth, the next woman assigned to that clinic was given a meter. Women in one group performed 7 self-monitored glucose determinations a day (fasting, preprandial, 2 hour postprandial, and at bedtime) (7 times daily monitoring group) and women in the other group were assessed weekly for fasting and 2-hour postprandial measurements during clinic visits and performed 4 self-monitored glucose determinations a day (fasting and 2 hours after breakfast, lunch, and dinner) (4 times daily monitoring group). Women in both groups were treated with either diet and insulin or diet alone. Diet was prescribed as 25 to 35 kilocalories per kilogram of body weight. Women who did not achieve glycaemic goals with diet alone were assigned to insulin therapy.	Length of stay in neonatal intensive care unit 4 times daily monitoring= 4.43 +/- 3 7 times daily monitoring= 2.77 +/- 2 p<0.0001 Neonatal hypoglycaemia 4 times daily monitoring= 263/1316 (20.0%) 7 times daily monitoring= 44/1145 (3.8%) p<0.0001 Shoulder dystocia 4 times daily monitoring= 18/1316 (1.4%) 7 times daily monitoring= 5/1145 (0.4%) p<0.0001 Stillbirth rate 4 times daily monitoring= 4/1000 7 times daily monitoring= 1/1000 p value not reported Neonatal death rate 4 times daily monitoring= 2/1000 7 times daily monitoring= 3/1000 p value not reported	A3 Groups were comparable at baseline, including all major confounding and prognostic factors – Yes B1 Comparison groups received the same care apart from the intervention(s) studied – Yes B2 Participants receiving care were kept 'blind' to treatment allocation – No B3 Individuals administering care were kept 'blind' to treatment allocation – No C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? – 69 women were lost to follow up (see 'other information' below) C2 b. Groups were comparable for treatment completion - Unclear C3 a. For how many

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	7 times daily monitoring= 48.9% p value not significant Gestational age at entry to diabetic program 4 times daily monitoring= 28 +/- 5 weeks 7 times daily monitoring= 27 +/- 6 weeks p value not significant Glucose screening result (mg/dL) 4 times daily monitoring= 182 +/- 47 7 times daily monitoring= 179 +/- 33 p value not significant Number of abnormal values on glucose tolerance test: 1 4 times daily monitoring= 33.1 7 times daily monitoring= 32.7 p value not significant 2 4 times daily monitoring= 37.9 7 times daily monitoring= 36.0 p value not significant 3 4 times daily monitoring=		All women were treated to attain the same mean blood glucose levels. Large for gestational age was defined as ≥90th percentile on the basis of growth standards developed for the population Hypoglycaemia was diagnosed if any two consecutive values of plasma glucose were ≤30 mg/dL (capillary heel blood). Outcomes were compared with χ2, Fisher's exact test, Student's t test, or analysis of variance		participants in each group were no outcome data available? - Unclear C3 b. Groups were comparable with respect to the availability of outcome data - Unclear D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - Unclear D5 Investigators were kept 'blind' to other important confounding and prognostic factors - Unclear Other information 69 women (2.7%) left the study because they gave birth at a different center (these women were not included in the analysis)

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	22.4 7 times daily monitoring= 23.1 p value not significant 4 4 times daily monitoring= 6.6 7 times daily monitoring= 8.2 p value not significant Oral glucose tolerance test (mg/dL): Fasting 4 times daily monitoring= 104 +/- 18 7 times daily monitoring= 102 +/- 21 p value not significant 1 hour 4 times daily monitoring= 199 +/- 30 7 times daily monitoring= 201 +/- 29 p value not significant 2 hour 4 times daily monitoring= 179 +/- 38 7 times daily monitoring= 178 +/- 31 p value not significant 3 hour 4 times daily monitoring= 136 +/- 40 7 times daily monitoring= 137 +/- 31 p value not significant				

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Inclusion criteria Women with gestational diabetes Exclusion criteria None reported				

A.9 HbA_{1c} monitoring during pregnancy

No evidence table

A.10 Ketone monitoring during pregnancy

No evidence table

A.11 What is the effectiveness of continuous glucose monitoring in pregnant women with diabetes compared with intermittent capillary blood glucose monitoring?

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation Kerssen,A., De Valk,H.W., Visser,G.H., Do HbA(1)c levels and the self-monitoring of blood glucose levels adequately reflect glycaemic	Sample size 43 women Characteristics Not reported	Interventions Continuous glucose monitoring with intermittent	Details The study was approved by the ethics committee of the University Medical Centre Utrecht, The Netherlands. All women gave written	Results Mean glucose level (mmol/l): 4 to 5 times a day group* Intermittent monitoring = 6.8 Continuous monitoring = 6.9	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
control during pregnancy in women with type 1 diabetes mellitus?, Diabetologia, 49, 25-28, 2006 Ref Id 252456 Country/ies where the study was carried out The Netherlands Study type Prospective within-subjects comparison Aim of the study To determine whether, in pregnant women with type 1 diabetes, HbA_{1c} levels within 1% above normal are appropriate or whether treatment should be aimed at normal HbA_{1c} levels, and to determine how many self-monitored blood glucose levels are needed each day to obtain an adequate image of glycaemic control. Study dates December 2001 to June 2004 Source of funding Supported by Novo Nordisk Farma BV, Alphen aan de Rijn, The Netherlands	Inclusion criteria None reported Exclusion criteria None reported	monitoring (n = 43)	informed consent to participate. Women were recruited from an obstetrical out-patient clinic. Women were asked to use continuous glucose monitoring once in each trimester of pregnancy, whilst continuing their regular self monitored blood glucose measurement (with a minimum of 4 self monitoring blood glucose measurements a day as this is the amount needed to calibrate the continuous glucose monitoring system). Women were asked to maintain their regular testing schedule for their self monitoring blood glucose measurements. All self monitored blood glucose measurements were performed using fingerstick measurement and the MediSense Precision Xtra glucose meter (Abbott, Bedford, MA, USA). HbA_{1c} levels were determined within 1 week after continuous glucose measurement. For 55% of the women, HbA_{1c} values	6 to 9 times a day group* Intermittent monitoring= 6.5 Continuous monitoring= 6.3 10 or more times a day group* Intermittent monitoring = 6.2 Continuous monitoring = 6.3 Hypoglycaemia episodes: 4 to 5 times a day group* Intermittent monitoring = 0.6** Continuous monitoring = 2.3** 6 to 9 times a day group* Intermittent monitoring = 1.2** Continuous monitoring = 2.5** 10 or more times a day group* Intermittent monitoring = 2.7** Continuous monitoring = 3.7** No adverse events were reported with the use of the continuous glucose monitoring system. It is not clearly reported in the paper what the denominators are. Self monitored blood glucose measurements were performed 4 or 5 times a day on 92 days, 6 to 9 times a day on 70 days, and 10 or more times a day on 23 days. * The number of measurement days that fulfilled the predetermined	to treatment groups was unrelated to potential confounding factors – N/A A2 Attempts were made within the design or analysis to balance the comparison groups for potential confounders – N/A A3 Groups were comparable at baseline, including all major confounding and prognostic factors – N/A B1 Comparison groups received the same care apart from the intervention(s) studied – N/A B2 Participants receiving care were kept 'blind' to treatment allocation – N/A B3 Individuals administering care were kept 'blind' to treatment allocation – N/A C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? - None C2 b. Groups were

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			were also obtrained 6 to 8 weeks after the continuous glucose monitoring measurement. HbA_{1c} values obtained 1 week or 6 to 8 weeks after the continuous glucose monitoring measurement were not significantly different. Glucose profiles measured with the continuous glucose monitoring system were only included if 288 glucose measurements were available per 24 hours (i.e. none were missing) and the following criteria were met: 1) at least four paired sensor glucose values and meter glucose readings per day, 2) correlation coefficient between sensor glucose values and meter blood glucose readings ≥ 0.79, and 3) average value of differences between sensor glucose values and meter glucose readings for a given day ≤ 28%. Hypoglycaemia was defined as a glucose level ≤ 3.9 mmol/l Measurement days were categorised into three groups depending on the number of daily self	requirements were 68 in the first trimester, 59 in the second trimester, and 58 in the third trimester. However, it is not clear how many women were in each group. **It is not clear whether this is a mean value for the group, for each woman, or for each day.	comparable for treatment completion - Yes C3 a. For how many participants in each group were no outcome data available? - None C3 b. Groups were comparable with respect to the availability of outcome data - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - No D5 Investigators were kept 'blind' to other important confounding and prognostic factors - No Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			monitoring blood glucose determinations: 4 or 5 determinations, 6 to 9 determinations, or 10 or more determinations.		
Full citation Kestila,Kirsimarja K., Ekblad,Ulla U., Ronnemaa,Tapani, Continuous glucose monitoring versus self-monitoring of blood glucose in the treatment of gestational diabetes mellitus, Diabetes Research and Clinical Practice, , 174-179, 2007 Ref Id 253163 Country/ies where the study was carried out Finland Study type Randomised trial Aim of the study To compare a continuous glucose monitoring system with self-monitoring of plasma glucose in determining whether women with gestational diabetes need antidiabetic drug treatment Study dates Not reported Source of funding	Sample size 73 women Characteristics Ethnicity: Finnish = 72/73 (99%) Indonesian = 1/73 (1%) Age (years): Intermittent group = 32.2 +/- 5.7 Continuous group = 32.6 +/- 4.7 p = 0.72 Primipara: Intermittent group = 20/37 (55.5%) Continuous group = 15/36 (41.7%) p = 0.15 BMI (kg/m2): Intermittent group = 26.1 +/- 3.3 Continuous group = 27.2 +/- 3.9 p = 0.18 Smokers:	Interventions Intermittent group = 37 women Continuous group = 36 women	Details The study was approved by the Turku University Hospital ethics committee. All women who participated gave written consent. Women were randomly allocated either to continuous glucose monitoring system (CGMS Medtronic MiniMed, Northridge, CA, USA) (continuous group) or self-monitoring of plasma glucose (intermittent group). The method of randomisation was not reported. Plasma glucose was measured with either Ascensia Elite meter (Bayer Corporation, Mishawaka, IN, USA), or Super Glucocard II meter (Arkray, Kyoto, Japan). All women came to the hospital for an interview and dietary counselling for low glycaemic index, low saturated fat eucaloric diet.	Results Spontaneous delivery: Intermittent group = 26/37 (70.3%) Continuous group = 25/36 (69.4%) p = 0.47 Assisted delivery: Intermittent group = 3/37 (8.1%) Continuous group = 3/36 (8.3%) p = 0.49 Caesarean section: Intermittent group = 8/37 (21.6%) Continuous group = 8/36 (22.2%) p = 0.47 Premature birth (< 37 gestational weeks): Intermittent group = 2/37 (5%) Continuous group = 2/36 (6%) p value not reported There were no births prior to 35 gestational weeks Gestational weeks at birth: Intermittent group = 39 + 5 +/- 1.3 Continuous group = 39 + 2 +/- 1.3 p = 0.22	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials A1 An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups) - unclear A2 There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation) - unclear A3 The groups were comparable at baseline, including all major confounding and prognostic factors - yes B1 The comparison groups received the same care apart from the intervention(s) studied - yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Turku University Central Hospital Research Fund, and The Foundation of Gynaecologists and Obstetricians in Finland supported the study	Intermittent group = 5 (13.5%) Continuous group = 4 (11.1%) p = 0.38 Hypertension: Intermittent group = 2 (5.4%) Continuous group = 4 (11.1%) p = 0.19 HbA_{1c} at start of study: Intermittent group = 5.3 +/- 0.3 Continuous group = 5.4 +/- 0.4 p = 0.15 Gestational weeks at birth: Intermittent group = 39 + 5 +/- 1.3 Continuous group = 39 + 2 +/- 1.3 p = 0.22 Inclusion criteria Women with gestational diabetes Women with singleton pregnancies Exclusion criteria None reported		All women were shown how to measure plasma glucose and asked to measure it at least 5 times a day (fasting plasma glucose, pre-prandial values, postprandial values at 90 minutes after main meals) as well as to keep a dietary and exercise diary on glucose measurement days. Women randomised for continuous glucose monitoring were also shown how to use the equipment. A minimum of 4 daily plasma glucose calibration values were used with the continuous glucose monitoring equipment. HbA_{1c} values were analysed using the Mann-Whitney test. It is not clear which method of statistical analysis was used for the other reported outcomes.	Macrosomia: Intermittent group = 3/37 (8.1%) Continuous group = 4/36 (11.1%) p = 0.33 Days per treated neonate in NICU: Intermittent group = 3.83 +/- 2.0 Continuous group = 3 +/- 1.3 p value not reported Neonates transferred to NICU: Intermittent group = 11/37 (30.8%) Continuous group = 7/36 (19.4%) p = 0.11 There were no perinatal deaths in either group No skin infections were observed where the electrodes were placed An average of 568 +/- 30 glucose measurements were recorded for each women using the continuous glucose monitoring system.	B2 Participants receiving care were kept 'blind' to treatment allocation - no B3 Individuals administering care were kept 'blind' to treatment allocation - no C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - yes C2 a. How many participants did not complete treatment in each group? – none C2 b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment) - yes C3 a. For how many participants in each group were no outcome data available? - none C3 b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					for whom outcome data were not available) - yes D1 The study had an appropriate length of follow-up - yes D2 The study used a precise definition of outcome - yes D3 A valid and reliable method was used to determine the outcome - yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - no D5 Investigators were kept 'blind' to other important confounding and prognostic factors - unclear Other information The women in this study were tested for gestational diabetes as they belonged to a high-risk group, due to: body mass index over 25 kg/m2, aged over 40 years, a previous child over 4500g, glucosuria during pregnancy, weight gain of more than 20kg during pregnancy, previous gestational diabetes, or suspectved foetal macrosomia in current pregnancy.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					The authors note that the study is not powered to detect any differences in obstetrical outcome between the two groups.
Full citation Murphy,H.R., Rayman,G., Lewis,K., Kelly,S., Johal,B., Duffield,K., Fowler,D., Campbell,P.J., Temple,R.C., Effectiveness of continuous glucose monitoring in pregnant women with diabetes: randomised clinical trial, BMJ, 337, a1680-, 2008 Ref Id 234219 Country/ies where the study was carried out UK Study type Randomised trial Aim of the study To determine the effectiveness of continuous glucose monitoring during pregnancy in women with type 1 and type 2 diabetes on maternal and neonatal outcomes Study dates September 2003 to 2006	Sample size 71 women Characteristics Type of diabetes: Type 1 = 46/71 (65%) Type 2 = 25/71 (35%) Mean age: Both groups = 31.3 +/- 6.1 years Intermittent group = 32.5 +/- 5.9 years Continuous group = 30.2 +/- 6.3 years p value not significant Diabetes type 1: Intermittent group = 18/33 (55%) Continuous group = 28/38 (74%) p value not reported Diabetes type 2: Intermittent group = 15/33 (45%) Continuous group =	Interventions Intermittent group = 33 women Continuous group = 38 women	Details The trial was conducted in two secondary care diabetic antenatal clinics in the UK. Women were approached consecutively and were included if they provided written informed consent and were willing to wear a continuous glucose monitor. 71 of 93 (76%) of women approached agreed to participate. Reasons for women not wishing to participate included not being interested in the study, social issues or problems with transport, work commitments, unwilling to wear the continuous glucose monitor, previous stillbirth, having young children, and being new to the area. No significant differences were found between women who participated or who declined in age, ethnicity, type or duration of diabetes, HbA_{1c} level or gestational age at booking, attendance at pre-	Results Vaginal birth: Intermittent group = 12/33 (39%) Continuous group = 11/38 (29%) p = 0.4 Elective caesarean: Intermittent group = 5/33 (20%) Continuous group = 16/38 (42%) p = 0.07 Emergency caesarean: Intermittent group = 13/33 (43%) Continuous group = 11/38 (29%) p = 0.3 All caesareans (elective and emergency): Intermittent group = 18/33 (55%) Continuous group = 27/38 (71%) p value not reported Pre-term delivery < 37 weeks: Intermittent group = 6/33 (19%) Continuous group = 6/38 (16%) p = 0.8	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials A1 An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups) - yes A2 There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation) - yes A3 The groups were comparable at baseline, including all major confounding and prognostic factors - unclear B1 The comparison

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Source of funding Funded by the Ipswich Diabetes Centre Charity Research Fund. One author received salary support from Diabetes UK. Study equipment was donated free of charge by Medtronic UK (6 CGMS Gold monitors and 300 sensors). The research was sponsored by Ipswich Hospital NHS Trust and was independent of all the study funders.	10/38 (26%) p value not reported Mean duration of diabetes: Both groups= 12.8 +/- 0.3 years Intermittent group = 10.0 +/- 8.8 years Continuous group = 15.2 +/- 11.0 years p = 0.03 Primiparous: Intermittent group = 11/33 (33%) Continuous group = 16/38 (42%) p value not reported Ethnicity White European: Intermittent group = 29/33 (88%) Continuous group = 34/38 (89%) p value not reported Asian: Intermittent group = 3/33 (9%) Continuous group = 3/38 (8%) p value not reported Other: Intermittent group= 1/33 (3%)		pregnancy care, and folic acid supplementation. Women were allocated to standard care (intermittant group) or standard care with the addition of a continuous glucose monitor (continuous group). Women were randomised using computer generated randomised numbers in blocks of 20, concealed in sealed envelopes. Women were provided with their group allocation by trained research nurses. Continuous glucose monitoring was offered supplementary to women's care for up to 7 days at intervals of 4 to 6 weeks between 8 and 32 weeks of gestation to reduce potentially greater discomfort in later pregnancy. The continuous glucose monitor (CGMS Gold Medtronic-MiniMed, Northridge, USA) measured glucose values every 10 seconds with an average value stored every 5 minutes, providing up to 288 measurements a day. The system was recallibrated each time a capillary glucose measurement was entered, and women were advised	HbA_{1c} 28 to 32 weeks' gestation: Intermittent group = 6.4% (SD 0.8) Continuous group = 6.1% (SD 0.6) p = 0.1 32 to 36 weeks' gestation: Intermittent group = 6.4% (SD 0.7) Continuous group = 5.8% (SD 0.6) p = 0.007 Early neonatal deaths: Intermittent group = 1/33 (3%) (singleton, 28 weeks) Continuous group = 1/39 (3%) (1 twin, at 34 weeks) p = 1.0 Macrosomia (≥ 90th centile): Intermittent group = 18/33 (60%) Continuous group = 13/39 (33%) p = 0.05 Extremely large for gestational age (≥ 97.7th centile): Intermittent group = 9/33 (30%) Continuous group = 5/39 (13%) p = 0.1 Admission to neonatal care unit: Intermittent group = 6/33 (19%) Continuous group = 9/39 (23%) p = 0.8 29/36 (80%) of the women wore the monitor at least once per trimester.	groups received the same care apart from the intervention(s) studied - yes B2 Participants receiving care were kept 'blind' to treatment allocation - no B3 Individuals administering care were kept 'blind' to treatment allocation - no C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - yes C2 a. How many participants did not complete treatment in each group? – 2 in the continuous group, 0 in the intermittent group C2 b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment) - unclear C3 a. For how many participants in each group were no outcome data available? - none C3 b. The groups were comparable with respect

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Continuous group= 1/38 (3%) p value not reported Mean body mass index (kg/m2): Both groups = 28.1 +/- 7.4 Intermittent group = 28.4 +/- 8.1 Continuous group = 27.9 +/- 7.0 p value not significant Mean HbA_{1c} value at booking: Both groups = 7.3 +/- 1.2% Intermittant group = 7.4 +/- 1.5% Continuous group = 7.2 +/- 0.9% p value not significant Mean gestational age at booking: Both groups = 9.2 +/- 2.7 weeks Intermittent group = 9.0 +/- 3.0 weeks Continuous group = 9.4 +/- 2.3 weeks p value not significant Pre-pregnancy care: Intermittent group=		to recallibrate the instrument at least 4 times a day. Trained research nurses with no clinical input implanted the sensors. Neither the participants nor the clinicians had access to the glucose measurements whilst the sensors were being used. Sensors were removed after 5 to 7 days unless they experienced pain, discomfort or technical problems. Women discussed the intermittent glucose monitoring data either with or without the continuous glucose monitoring data (depending on which group the women were allocated to) with a diabetes specialist nurse. Women were asked to note down the likely causes of unusual patterns of hypoglycaemia or hyperglycaemia, and to suggest possible solutions, including changes to diet, activity and insulin dose. In the first meeting this was done in conjunction with the research team, but thereafter was done with the woman's support person. The suggested change to diet, activity, and insulin dose were then discussed with the obstetric diabetes team based on	Mean number of periods of continuous glucose monitoring in the 36 women whose pregnancies did not end prematurely = 4.2 (range 0 to 8). The continuous glucose monitor was 'generally well tolerated'. There were no skin infections, although mild erythema and inflammation were reported around the insertion point in some women. 1 woman experienced pain after insertion of the sensor and withdrew from the study. 1 woman declined to participate after the first continuous glucose profile had been downloaded. Some women reported a reduced use of the continuous glucose monitor, for the following reasons: discomfort, transport, and difficulties with bathing. 3 infants in each group were excluded from the analysis of birthweight centile as a result of miscarriage in the first trimester, neonatal death, a major malformation. There were 2 sets of living twins, plus 1 single surviving twin, resulting in 5 healthy babies resulting from twin pregnancies (all in the continuous group). The analyses for birthweight centile were done both with twins (using the appropriate centile reference range for twins) and without twins,	to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available) - yes D1 The study had an appropriate length of follow-up - yes D2 The study used a precise definition of outcome - yes D3 A valid and reliable method was used to determine the outcome - yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - no D5 Investigators were kept 'blind' to other important confounding and prognostic factors - unclear Other information A power calculation conducted by the authors stated that a sample size of 70% would give 80% power to detect a 40% reduction in macrosomia at p = 0.05, based on a macrosomia rate of 60%. A sample size of 70 would give a

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	18/33 (55%) Continuous group= 24/38 (63%) p value not reported Folic acid at booking: Intermittent group = 27/33 (82%) Continuous group = 33/38 (87%) p value not reported Microvascular complication: Intermittent group = 3/33 (10%) Continuous group = 7/38 (18%) p value not reported Smoker: Intermittent group = 4/33 (12%) Continuous group = 5/38 (13%) p value not reported Information on maternal characteristics was obtained from hospital maternity records. Inclusion criteria Women aged 16 to 45 years old Women with type 1 or type 2 diabetes		the intermittent data alone or in conjunction with the continuous data. Information on HbA_{1c} levels were obtained from hospital maternity records. Women were asked to measure blood glucose levels at least 7 times a day - before meals, one hour after meals, and two hours after meals. Women were seen every 2 to 4 weeks for up to 28 weeks, fortnightly until 32 weeks, and weekly thereafter, with assessments of fetal growth at 28, 32, and 36 weeks. HbA_{1c} levels were measured once every 4 weeks. HbA_{1c} values were compared using t tests Birthweight centiles were compared using Wilcoxon rank sum test Macrosomia was compared using Fisher exact tests	and there was no change to the significance of the results.	50% reduction in risk at 95% power.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Exclusion criteria Severe medical or psychological comorbidity				
Full citation Secher,Anna L., Ringholm,Lene, Andersen,Henrik U., Damm,Peter, Mathiesen,Elisabeth R., The Effect of Real-Time Continuous Glucose Monitoring in Pregnant Women With Diabetes: A randomized controlled trial, Diabetes Care, E-Publish ahead of print, -, 2013 Ref Id 259104 Country/ies where the study was carried out Denmark Study type Randomised controlled trial Aim of the study To determine whether continuous glucose monitoring is beneficial to women with diabetes during pregnancy Study dates February 2009 to February 2011	Sample size 154 women Characteristics Age (years, median): Continuous monitoring= 32 (range 21 to 42) Intermittent monitoring= 31 (range 19 to 43) p= 0.88 Pregestational BMI (kg/m2, median): Continuous monitoring= 25.1 (range 18.6 to 52.7) Intermittent monitoring= 24.7 (range 18.4 to 48.2) p= 0.69 Type 1 diabetes= 123 (80%) Type 2 diabetes= 31 (20%) 27 (22%) women with type 1 diabetes were on insulin pump therapy 30 (97%) women with type 2 diabetes received insulin therapy during pregnancy	Interventions Continuous monitoring (n= 79) Intermittent monitoring (n= 75)	Details The research protocol was approved by the Danish National Committee on Biomedical Research Ethics and the Danish Data Protection Agency. Women who participated gave written informed consent. All Danish speaking pregnant women with diabetes prior to pregnancy with one living intrauterine fetus who were referred to the Center for Pregnant Women with Diabetes Rigshospitalet prior to 14 weeks completed gestation were invited to take part in the study. Women who had more than one pregnancy during the study period (n= 4) were only offered inclusion in the study at referral for their first pregnancy. Women were randomised using a computer-generated randomisation program (no further details given). Treatment allocation was concealed using an automated telephone allocation	Results Caesarean section Continuous monitoring= 28/79 (37%) Intermittent monitoring= 33/75 (45%) p= 0.30 Pre-term birth Continuous monitoring= 16/79 (21%) Intermittent monitoring= 12/75 (16%) p= 0.47 HbA_{1c} (%, median) (76 women in continuous group, 73 in intermittent group): 8 weeks Continuous monitoring= 6.6 (range 5.3 to 10.0) Intermittent monitoring= 6.8 (range 5.3 to 10.7) p= 0.72 33 weeks Continuous monitoring= 6.1 (range 5.1 to 7.8) Intermittent monitoring= 6.1 (range 4.8 to 8.2) p= 0.39	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials A1 An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups) - yes A2 There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation) - yes A3 The groups were comparable at baseline, including all major confounding and prognostic factors - yes B1 The comparison groups received the same care apart from the intervention(s)

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Source of funding One author received financial support from the European Foundation for the Study of Diabetes and LifeScan, Rigshospitalet's Research Foundation, the Capital Region of Denmark, the Medical Faculty Foundation of Copenhagen University, Aase and Ejnar Danielsen Foundation, and Master Joiner Sophus Jacobsen and his wife Astrid Jacobsen's Foundation. One author holds stock in Novo Nordisk. One author received financial support from the Novo Nordisk Foundation. Medtronic supplied the study with real-time continuous glucose monitors and links and glucose sensors were offered at a reduced price.	During the study period, 30 women received antihypertensive medication, 8 women received antidepressive medication, and 32 women were treated for thyroid dysfunction. Duration of diabetes (years, median): Continuous monitoring= 10 (range 1 to 37) Intermittent monitoring= 12 (range 1 to 38) p= 0.38 HbA_{1c} at baseline (%, median): Continuous monitoring= 6.6 (range 5.3 to 10.0) Intermittent monitoring= 6.8 (range 5.3 to 10.7) p= 0.67 Diabetic retinopathy: Continuous monitoring= 28 (35%) Intermittent monitoring= 32 (44%) p= 0.29 Elevated urine albumin excretion (albumin-to-creatinine ratio ≥30mg/mmol in a random urine sample):		service provided by an external organisation. Women were stratified according to their type of diabetes. Women in both groups followed a routine pregnancy care program. All women had a dietitian appointment at their first pregnancy visit. Women were given weight targets based on their BMI. Women in the other group were offered continuous glucose monitoring at 8, 12, 21, 27 and 33 weeks for 6 days (continuous monitoring group). Some women were only willing to use continuous monitoring for 3 days per monitoring period, which was allowed. The majority of women had the sensor inserted in the abdominal skin, although later in pregnancy some women had it inserted in their upper arm. Women were taught how to use the continuous glucose monitors and were requested to continue taking intermittent measurements. Therapeutic adjustments to diet, exercise, and insulin doses were primarily based on intermittent monitoring values. The women in the one	36 weeks Continuous monitoring= 6.0 (range 5.1 to 7.7) Intermittent monitoring= 6.1 (range 4.7 to 8.4) p= 0.63 At least 1 severe hypoglycaemic event: All women Continuous monitoring= 13/79 (16%) Intermittent monitoring= 12/75 (16%) p= 0.91 Women with type 1 diabetes using continuous montioring per protocol Continuous monitoring= 4/38 (11%) Intermittent monitoring= 11/59 (19%) p= 0.28 By type of diabetes (across both study arms) Type 1= 19/123 (16%) Type 2= 5/31 (17%) p value not significant (actual value not reported) Miscarriage Continuous monitoring= 3/79 (4%) Intermittent monitoring= 2/75 (3%) p value not reported Large for gestational age infant	studied - yes B2 Participants receiving care were kept 'blind' to treatment allocation - no B3 Individuals administering care were kept 'blind' to treatment allocation - no C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - yes C2 a. How many participants did not complete treatment in each group? – none C2 b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment) - yes C3 a. For how many participants in each group were no outcome data available? - none C3 b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Continuous monitoring= 5 (6%) Intermittent monitoring= 2 (3%) p= 0.44 Smoker: Continuous monitoring= 6 (8%) Intermittent monitoring= 9 (12%) p= 0.34 Inclusion criteria Pregnant women with pre-existing diabetes Exclusion criteria Use of continuous monitoring at time of recruitment into the study (n= 7) Severe mental or psychiatric barriers (n= 4) Diabetic retinopathy (n= 3) Severe concurrent comorbidity (1= severe psoriasis, 2= previous gastric bypass surgery)		group were recommended to monitor their plasma glucose measurements 8 times daily (before and 90 minutes after each main meal, before bed, and at 3am) for 6 days, at 8, 12, 21, 27, and 33 weeks (intermittent monitoring group). Diet and insulin doses were adjusted by the women every third day themselves, and with an experienced diabetologist every two weeks. A power calculation found that 45 women were needed in each arm (based on assumption of prevalence of 50% large for gestational age babies in study population, and that continuous monitoring could reduce this to 20%). Characteristics of the groups were compared using the Fisher exact test or $\chi 2$ for dichotomous variables and t test or Mann-Whitney for continuous variables. A $p<0.05$ was considered significant. Analyses were performed on an intention-to-treat basis, including 154 women at baseline and excluding women with	Continuous monitoring= 34/79 (45%) Intermittent monitoring= 25/75 (34%) p= 0.19 1 incidence of perinatal death in a woman with type 2 diabetes due to severe should dystocia, however, it is not clear which treatment group this woman was in Continuous monitoring was generally well tolerated without severe side effects. 49 (64%) of women used continuous monitoring per protocol (i.e. during the weeks requested by the study authors), and 5 (7%) of women used it at least 60% of the time throughout pregnancy.	groups in terms of those for whom outcome data were not available) - yes D1 The study had an appropriate length of follow-up - yes D2 The study used a precise definition of outcome - yes D3 A valid and reliable method was used to determine the outcome - yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - no D5 Investigators were kept 'blind' to other important confounding and prognostic factors - unclear Other information 47 women who were eligible did not participate - they were similar to the included women for all baseline characteristics except they had a slightly shorter duration of diabetes (actual data not reported). The main reason women declined to participate was the possibility of being given continuous glucose

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			miscarriages (n=5) from the outcome data Mild hypoglycaemia was defined as "events familiar to the patient as hypoglycaemia and managed by the patient" Severe hypoglycaemia was defined as "self-reported events with symptoms of hypoglycaemia requiring help from another person to actively administer oral carbohydrate or injection of glucose or glucagon in order to restore normal blood glucose level" Large for gestational age was defined as ≥90th percentile adjusted for sex and gestational age		monitoring
Full citation Yogev,Y., Chen,R., Ben-Haroush,A., Phillip,M., Jovanovic,L., Hod,M., Continuous glucose monitoring for the evaluation of gravid women with type 1 diabetes mellitus, Obstetrics and Gynecology, 101, 633-638, 2003 Ref Id 213994 Country/ies where the study was carried out Israel	Sample size 34 women Characteristics All women had type 1 diabetes prior to the onset of pregnancy Gestational age: Range 16 to 32 weeks All women were being treated with insulin and had individualised counselling from a dietitian	Interventions Continuous glucose monitoring with intermittant monitoring (n = 34)	Details The study protocol was approved by the local ethics committee. Women were recruited consecutively during a routine clinical visit to the Diabetes in Pregnancy Centre of the Perinatal Division Unit, Rabin Medical Centre. Women were included if they gave consent to participate after an explanation of the study	Results Mean glucose level (mg/dl): Intermittent monitoring = 101 +/- 13 Continuous monitoring = 121 +/- 13 p = 0.02 No adverse events associated with the use of continuous glucose monitoring were reported. None of the women experienced irritation or infection at the insertion site. Women reported high satisfaction using the device concerning future benefits of continual monitoring	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies A1 Method of allocation to treatment groups was unrelated to potential confounding factors – N/A A2 Attempts were made within the design or analysis to balance the comparison groups for potential confounders –

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Study type Prospective within-subjects comparison Aim of the study To compare the daily glycemic profile as shown by continuous and intermittent blood glucose monitoring in pregnant women with type 1 diabetes. The study also examined whether treatment strategy protocols based on the two monitoring methods differed. Study dates November 2001 to March 2002 Source of funding None reported	Mean age: 26 +/- 4.7 years (range 21 to 36 years) Mean gestational age (?at recruitment): 25 +/- 6.2 weeks (range 16 to 21 weeks) Mean gravidity: 2.4 +/- 1.1 Mean parity: 1.2 +/- 0.9 Mean BMI: 26.2 +/- 4.7 kg/m2 Mean HbA_{1c} level: 6.1 +/- 1.2% (normal range 4.5 to 5.7%) Inclusion criteria Not reported Exclusion criteria Not reported		(83% of the women approached were included). A MiniMed continuous glucose monitoring system (MiniMed, Sylmar, CA) was used in all women for 3 days. The same nurse placed all of the continuous glucose monitoring sensors. Glucose measurements are taken by the system every 10 seconds, which stores an average value every 5 minutes, giving a total of 288 measurements a day. The women were unaware of the sensor measurements during the monitoring period, but were trained how to code the time of food intake, insulin injections, exercise periods, and symptomatic hypoglycaemia into the monitor. Women were asked to wear the continuous glucose monitoring device for 72 consecutive hours whilst also performing fingerstick capillary glucose measurements in the morning after overnight fasting and 2 hours after meals (6 to 8 times a day) using a glucometer (Ames Glucometer Elite, Bayer	All women completed the 3 day study An average of 780 +/- 54 glucose measurements was recorded for each woman with continuous glucose monitoring	N/A A3 Groups were comparable at baseline, including all major confounding and prognostic factors – N/A B1 Comparison groups received the same care apart from the intervention(s) studied – N/A B2 Participants receiving care were kept 'blind' to treatment allocation – N/A B3 Individuals administering care were kept 'blind' to treatment allocation – N/A C1 All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes C2 a. How many participants did not complete treatment in each group? - None C2 b. Groups were comparable for treatment completion - Yes C3 a. For how many participants in each group were no outcome data available? - None C3 b. Groups were comparable with respect

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			Corp., Elkhart, IN) and self-coding the data into the monitor. Quality control measures of glucose levels were taken from the meter and sensor at the time of connection to the continuous glucose monitoring system and at study completion. Data collected from self-blood glucose monitoring and continuous glucose monitoring were evaluated separately by one experienced clinician. A hypoglycaemic event was defined as a greater than 30 minute asymptomatic reading below 50 mg/dl or symptomatic hypoglycaemia detected by meter or monitoring records. Data were analysed using paired t-tests.		to the availability of outcome data - Yes D1 The study had an appropriate length of follow-up - Yes D2 The study used a precise definition of outcome - Yes D3 A valid and reliable method was used to determine the outcome - Yes D4 Investigators were kept 'blind' to participants' exposure to the intervention - No D5 Investigators were kept 'blind' to other important confounding and prognostic factors - No Other information

A.12 Screening for gestational diabetes in the first trimester

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation Bito,T., Nyari,T., Kovacs,L., Pal,A., Oral glucose tolerance testing at gestational weeks < or =16 could predict or exclude subsequent gestational diabetes mellitus during the current pregnancy in high risk group, European Journal of Obstetrics, Gynecology, and Reproductive Biology, 121, 51-55, 2005 Ref Id 152996 Country/ies where the study was carried out Hungary Study type Prospective cohort study Aim of the study To determine possible upper	Sample size 163 women at 16 gestational weeks or less were enrolled in the study Characteristics Patient characteristics are not presented for women diagnosed with gestational diabetes < 16 gestational weeks (these women were excluded from the study) Inclusion Criteria All pregnant women who did not have a previous history of gestational diabetes or any history of alteration of carbohydrate metabolism, but who displayed one or more risk factors for gestational diabetes and who were referred to the specialist outpatient department. The risk factors were: family history of type 2 diabetes, history of a large neonate (≥ 4000g), history of an adverse perinatal outcomes (missed abortion, malformation, polyhydramnios, stillbirth or preterm delivery), obesity (pre-pregnant BMI ≥ 30m2), age ≥ 35 years or glycosuria. Exclusion Criteria Women who were diagnosed as having gestational diabetes by OGTT at < 16 gestational weeks were excluded from the study	Tests Index test: No index test was used Reference standard: 2 hour 75g OGTT performed at 3 time periods: ≤ gestational week 16, gestational weeks 24-28 and gestational weeks 32-34 Diagnostic criteria: WHO 1999 thresholds for gestational diabetes - fasting plasma glucose value (FPG) ≥ 7 mmol/l and/or 2h postload plasma glucose value (2h PPG) ≥ 7.8 mmol/l	Methods For OGTT: Women were instructed to consume at least 150g of carbohydrate each day for 3 days and then to adhere to a 10-12 hour overnight fast the day before the OGTT. Venous plasma samples were collected at fasting and 2 hours after ingestion of 75g glucose solution over a 5 minute period. Glucose levels were determined by the glucose oxidase-peroxidase (GOD-POD) colorimetric method on sodium fluoride-mediated blood. The interassay and the interassay coefficient of variation were <2%.	Results Incidence of gestational diabetes Incidence of gestational diabetes at ≤ gestational week 16 = 8/163 (4.91%)* Incidence of gestational diabetes at ≤ week 16 / Incidence of gestational diabetes by gestational week 28 = 8/40 (20%)* Incidence of gestational diabetes at ≤ week 16 / Incidence of gestational diabetes by gestational week 34 = 8/88 (9.1%)*	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: There was no index test 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: There was no index test 7) Was the reference standard independent of

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
and lower cut-off values for the oral glucose tolerance test (OGTT) at or before gestational week 16 to predict subsequent onset of gestational diabetes in a high risk population and to assess the proportion of the group that would not require further OGTTs if these were applied Study dates 1 January 2001 to 30 September 2002 Source of funding Not stated					the index test i.e. the index test did not form part of the reference standard: There was no index test 8) Was the execution of the index test described in sufficient detail to permit its replication: There was no index test 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: There was no index test 11) Were the reference standard results interpreted without knowledge of the results of the index test: There was no index test 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: There were none

<table>
<tr><th>Bibliographic details</th><th>Participants</th><th>Tests</th><th>Methods</th><th>Outcomes and results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>14) Were withdrawals explained: There were none

Other information
* Calculated by NCC-WCH</td></tr>
<tr><td>Full citation
Kuti,M.A., Abbiyesuku,F.M., Akinlade,K.S., Akinosun,O.M., Adedapo,K.S., Adeleye,J.O., Adesina,O.A., Oral glucose tolerance testing outcomes among women at high risk for gestational diabetes mellitus, Journal of Clinical Pathology, 64, 718-721, 2011
Ref Id
153427
Country/ies where the study was carried out
Nigeria
Study type
Retrospective cohort study
Aim of the study</td><td>Sample size
765 pregnant women of whom 69 (9%) presented in, and had data available for, the first trimester

Characteristics
<table><tr><th></th><th>All</th><th>First trimester</th><th>Second trimester</th><th>Third trimester</th></tr><tr><td>Number of subjects</td><td>765</td><td>69</td><td>276</td><td>420</td></tr><tr><td>Age, years (mean, SD)</td><td>32.3 (4.4)</td><td>31.8 (4.1)</td><td>32.4 (4.5)</td><td>32.4 (4.4)</td></tr><tr><td>Positive family history of diabetes, n (%)</td><td>155 (20.3)</td><td>14 (20.3)</td><td>62 (22.5)</td><td>79 (18.8)</td></tr><tr><td>History of gestational diabetes, n (%)</td><td>14 (1.8)</td><td>2 (2.9)</td><td>6 (2.2)</td><td>6 (1.4)</td></tr></table>
Inclusion Criteria
Pregnant women referred to the Metabolic Research Unit (MRU) of University College</td><td>Tests
Index test: No index test was used
Reference standard: 2 hour 75g oral glucose tolerance test
Diagnostic criteria: WHO 1999 thresholds for gestational diabetes - fasting plasma glucose value (FPG) ≥ 7 mmol/l and/or 2h postload plasma glucose value ≥ 7.8 mmol/l</td><td>Methods
The records of all women referred between June 2007 and July 2009 were reviewed.
For OGTT: Following an overnight fast, two blood samples were taken before and 2h after a 75g of glucose load was administered orally. A diagnosis of gestational diabetes was made in accordance with the 1999 WHO guidelines. No details regarding standards of laboratory techniques are reported.</td><td>Results
Incidence of gestational diabetes

Incidence of gestational diabetes in the first trimester = 12/69 (17.4%)*

Incidence of gestational diabetes in the first trimester/ Incidence of gestational diabetes by end of second trimester = 12/47 (25.5%)*

Incidence of gestational diabetes in the first trimester/ Incidence of gestational diabetes by gestational week 40 = 12/106 (11.3%)*

* Calculated by NCC-WCH</td><td>Limitations
NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy
1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes
2) Were selection criteria clearly described: Yes
3) Was the reference standard likely to classify the target condition correctly: Yes
4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: There was no index test
5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The</td></tr>
</table>

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
To determine the prevalence and relationships with known risk factors of gestational diabetes at University College Hospital, Ibadan Study dates June 2007 to July 2009 Source of funding Not stated	Hospital, Ibadan for an oral glucose tolerance test (OGTT). Referrals were made for women at high risk of gestational diabetes based on a history of fetal macrosomia, maternal obesity, previous intrauterine fetal death, first degree relative with diabetes, glycosuria and history of gestational diabetes in a previous pregnancy. Exclusion Criteria Not stated				whole sample 6) Did participants receive the same reference standard regardless of the index test result: There was no index test 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: There was no index test 8) Was the execution of the index test described in sufficient detail to permit its replication: There was no index test 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: There was no index test 11) Were the reference standard results interpreted without knowledge of the results of the index test: There was no index test 12) Were the same clinical data available when the test results

<table>
<tr><th>Bibliographic details</th><th>Participants</th><th>Tests</th><th>Methods</th><th>Outcomes and results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>were interpreted as would be available when the test is used in practice: Yes
13) Were uninterpretable, indeterminate or intermediate test results reported: There were none
14) Were withdrawals explained: There were none

Other information</td></tr>
<tr><td>Full citation
Agarwal,M.M., Dhatt,G.S., Punnose,J., Zayed,R., Gestational diabetes: fasting and postprandial glucose as first prenatal screening tests in a high-risk population, Journal of Reproductive Medicine, 52, 299-305, 2007
Ref Id
153968
Country/ies where the study</td><td>Sample size
760 women who attended the antenatal clinic at Al Ain Hospital during the 12 month study period of whom 52 were unable to complete the oral glucose tolerance test (OGTT).
Therefore the total sample was 708 women included in the study (93.2%).

Characteristics<table><tr><th>Variable</th><th>Women without gestational diabetes</th><th>Women with gestational diabetes</th><th>p-value</th></tr><tr><td>Age (year) Mean±SD</td><td>27.9 ± 5.5</td><td>28.8 ± 5.5</td><td>0.09</td></tr><tr><td>Age (year) Median</td><td>27</td><td>28</td><td></td></tr><tr><td>Age (year) range</td><td>16-44</td><td>19-48</td><td></td></tr><tr><td>Gestational age</td><td>10.6 ± 2.5</td><td>10.4 ± 2.5</td><td>0.41</td></tr></table></td><td>Tests
Index test: Fasting plasma glucose (FPG)

Reference standard: 2 hour 75g OGTT.

Diagnostic criteria: WHO 1999 thresholds for gestational diabetes - FPG valule ≥ 126mg/dl (7.0 mmol/l) and/or 2h postload plasma glucose value</td><td>Methods
A universal screening strategy was used.

FPG and postprandial glucose (PPG) were tested at the first antenatal visit, usually in the first trimester. The FPG sample was collected after a 8-10 hour fast. A 2 hour 75g OGTT was performed within 2 weeks when the value of the FPG or PPG was ≥ 95mg/dl</td><td>Results
Incidence of gestational diabetes

In total, 184/708 (25.9%) women were diagnosed as having gestational diabetes
176/184 were diagnosed based on 2hr PPG ≥ 140mg/dl (7.8mmol/l)
8/184 women were diagnosed based on FPG ≥ 126mg/dl (7.0mmol/l)

79/184 (42.9%) were diagnosed as having gestational diabetes in first trimester (up to 18 gestational weeks)
105/184 (57.1%) were diagnosed in the second trimester (24-28 gestational weeks)

Diagnostic test accuracy of FPG index test at different</td><td>Limitations
NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy
1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes
2) Were selection criteria clearly described: Yes
3) Was the reference standard likely to classify the target condition correctly: Yes
4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target</td></tr>
</table>

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
was carried out United Arab Emirates Study type Prospective cohort study Aim of the study To determine the value of fasting plasma glucose and 2 hour postprandial plasma glucose as screening tests for gestational diabetes when performed at the first antenatal visit Study dates 1 September 2003 to 31 August 2004 Source of funding Not stated. Protocol was approved by the Research and Ethics Committee of the Faculty of Medicine and Health	(week) Mean±SD Gestational age (week) Median: 10 / 10 Gestational age (week) Range: 5 - 18 / 5-18 Fasting glucose (mg/dl) Mean ± SD: 89.8 ± 9.0 / 93.7 ± 13.1 / 0.001 Postprandial glucose (mg/dl) Mean ± SD: 98 ± 18.5 / 115 ± 24.9 / 0.001 BMI Mean ± SD: 26.5 ± 5.6 / 28.8 ± 7.1 / 0.001 Inclusion Criteria Women attending the antenatal clinic Exclusion Criteria None stated	≥ 140mg/dl (7.8mmol/l)	(5.3mmol/l) or ≥ 140mg/dl (7.8mmol/l) respectively. For the OGTT, venous plasma samples were collected for fasting (after a 12 hour overnight fast), and for 1 hour and 2 hour post glucose load. All women who tested negative on either screening test underwent a second diagnostic 2 hour 75g OGTT at 24-28 weeks gestation. The laboratory met the standards for both internal and external quality assurance for glucose.	thresholds in the first trimester compared with reference standard 2 hour OGTT interpreted using WHO 1999 criteria thresholds (FPG ≥ 7.0 or 2 hour PG ≥ 7.8 mmol/l) in the first trimester at FPG test threshold of 3.89mmol/l (70mg/dl) TP: 183* FN: 1* FP: 520* TN: 4* Sensitivity, % (95% CI): 99.5 (98.1 to 100)* Specificity, % (95% CI): 0.8 (0.3 to 0.9)* LR (95% CI): 1.00 (0.98 to 1.01)* LR- (95% CI): 0.71 (0.03 to 6.65)* at FPG test threshold of 4.17mmol/l (75mg/dl) TP: 181* FN: 3* FP: 505* TN: 19* Sensitivity, % (95% CI): 98.4 (95.8 to 99.6)* Specificity, % (95% CI): 3.6 (2.7 to 4.0)* LR (95% CI): 1.02 (0.98 to 1.04)* LR- (95% CI): 0.45 (0.11 to 1.57)* at FPG test threshold of 4.44mmol/l (80mg/dl) TP: 173* FN: 11* FP: 463* TN: 61* Sensitivity, % (95% CI): 94.0 (90.0 to 96.7)* Specificity, % (95% CI): 11.6 (10.2 to 12.6)* LR (95% CI): 1.06 (1.00 to 1.11)* LR- (95% CI): 0.51 (0.26 to 0.98)* at FPG test threshold of	condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: Yes. Index test was a FPG test that was not performed as part of the 2hr 75g OGTT 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Sciences, UAE University				4.72mmol/l (85mg/dl) TP: 147* FN: 37* FP: 380* TN: 144* Sensitivity, % (95% CI): 79.9 (74.2 to 84.9)* Specificity, % (95% CI): 27.5 (25.5 to 29.2)* LR (95% CI): 1.10 (1.00 to 1.20)* LR- (95% CI): 0.73 (0.52 to 1.01)* at FPG test threshold of 5.00mmol/l (90mg/dl) TP: 112* FN: 72* FP: 265* TN: 259* Sensitivity, % (95% CI): 60.9 (54.4 to 67.1)* Specificity, % (95% CI): 49.4 (47.2 to 51.6)* LR (95% CI): 1.20 (1.03 to 1.39)* LR- (95% CI): 0.79 (0.64 to 0.97)* at FPG test threshold of 5.28mmol/l (95mg/dl) TP: 72* FN: 112* FP: 165* TN: 359* Sensitivity, % (95% CI): 39.1 (33.0 to 45.4)* Specificity, % (95% CI): 68.5 (66.4 to 70.7)* LR (95% CI): 1.24 (0.98 to 1.55)* LR- (95% CI): 0.89 (0.77 to 1.01)* at FPG test threshold of 5.56mmol/l (100mg/dl) TP: 40* FN: 144* FP: 65* TN: 459* Sensitivity, % (95% CI): 21.7 (16.9 to 26.9)* Specificity, % (95% CI): 87.6	interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: No 14) Were withdrawals explained: Yes Other information

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				(85.9 to 89.4)* LR (95% CI): 1.75 (1.20 to 2.54)* LR- (95% CI): 0.89 (0.82 to 0.97)* at FPG test threshold of 5.83mmol/l (105mg/dl) TP: 21* FN: 163* FP: 28* TN: 496* Sensitivity, % (95% CI): 11.4 (7.9 to 15.2)* Specificity, % (95% CI): 94.7 (93.4 to 96.0)* LR (95% CI): 2.14 (1.20 to 3.79)* LR- (95% CI): 0.94 (0.88 to 0.99)* at FPG test threshold of 6.11mmol/l (110mg/dl) TP: 15* FN: 169* FP: 23* TN: 685* Sensitivity, % (95% CI): 8.2 (5.4 to 10.3)* Specificity, % (95% CI): 98.5 (97.5 to 99.2)* LR (95% CI): 5.34 (2.17 to 13.59)* LR- (95% CI): 0.93 (0.90 to 0.97)* * Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html	
Full citation Church,D., Halsall,D., Meek,C., Parker,R.A., Murphy,H.R., Simmons,D.,	Sample size Records were available for 26,369 live births although corresponding maternal data could not be matched for 506 cases. Characteristics are presented for 25,789 patients. 17,852 records included RBG test data.	Tests Index test: A screening RBG performed at the antenatal	Methods All women received venous plasma RBG measurement at antenatal booking as part of a	Results 17,852 records included RBG test data 3320*/17,852 (18.6%) women had RBG > 7.0mmol/l 3007 women had an OGTT	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of

<table>
<tr><th>Bibliographic details</th><th>Participants</th><th>Tests</th><th>Methods</th><th>Outcomes and results</th><th>Comments</th></tr>
<tr>
<td>Random blood glucose measurement at antenatal booking to screen for overt diabetes in pregnancy: a retrospective study, Diabetes Care, 34, 2217-2219, 2011
Ref Id
181105
Country/ies where the study was carried out
England
Study type
Retrospective cohort study
Aim of the study
To test the usefulness of a random venous blood glucose (RBG) taken at the booking appointment to detect overt diabetes in pregnancy (ODIP).

Study dates
Maternal and neonatal birth data from 2004-</td>
<td>Characteristics
Characteristics for women included and excluded from the study (n = 25,789)
<table>
<tr><td>Median (range) or Number (percentage)</td><td>Included patients n = 17,852</td><td>Excluded patients n = 7937</td></tr>
<tr><td>Maternal age years at birth</td><td>31 (13 - 54) n = 17,852</td><td>31 (15 - 49) n = 7936</td></tr>
<tr><td>Maternal BMI pre-pregnancy</td><td>24.0 (15.0 - 65.0) n = 15,611</td><td>23.0 (14.7 - 72.0) n = 6244</td></tr>
<tr><td>Parity:
Primiparous
Multiparous</td><td>6749 (37.9)
11,077 (62.1)
n = 17,826</td><td>3234 (41.1)
4628 (58.9)
n = 7862</td></tr>
<tr><td>Delivery method:
Spontaneous vaginal delivery
Elective CS
Emergency CS
Instrumental
Breech</td><td>10,397 (58.3)
2272 (12.7)
2773 (15.5)
2333 (13.1)
71 (4.0)
n = 17,846</td><td>4998 (63.9)
600 (7.7)
1192 (15.2)
986 (12.6)
48 (0.6)
n = 7824</td></tr>
<tr><td>Estimated Gestational age at birth:
< 32 weeks
33-41 weeks
> 42 weeks</td><td>263 (1.5)
17,022 (95.4)
566 (3.2)
n = 17,852</td><td>373 (4.7)
7256 (91.4)
308 (3.9)
n = 7937</td></tr>
<tr><td>Birth weight - g</td><td>3425 (340-5570) n = 17,846</td><td>3420 (50-5680) n = 7843</td></tr>
</table>
</td>
<td>booking appointment but defined as an RBG requested between 0 and 20 gestational weeks. If more than one RBG was identified for a woman, the highest value was used.

Reference test: A 75g oral glucose tolerance test (OGTT) using either venous or capillary sampling, performed at any time during gestation.

Diagnostic criteria: WHO 1999 thresholds for diabetes - fasting plasma glucose value (FPG) ≥ 7 mmol/l and/or 2h postload plasma</td>
<td>universal screening program. Women with a booking RBG >7.0 mmol/l or with a previous history of gestational diabetes were offered a 75g OGTT (venous or capillary sampling). Women diagnosed as not having gestational diabetes were screened again using a 50g oral glucose challenge test (GCT) at 26–28 weeks. Those with a GCT result > 7.7 mmol/l were offered an OGTT. OGTTs were also offered to women where it was clinically indicated (for example macrosomia).

Samples were collected using standard fluoride-containing tubes and analyzed in the hospital laboratory using a</td>
<td>during their pregnancy
87 women had RBG ≥ 11.1mmol/l and had an OGTT performed
26*/87 (30%) women had RBG ≥ 11.1mmol/l, had an OGTT performed and had diagnosed ODIP
67 women had a RBG and an OGTT performed and had diagnosed ODIP
12 women had RBG ≥ 11.1mmol/l and did not have an OGTT

Three analyses were performed to produce receiver operating curves (ROCs):

1) Using all 17,852 RBG data and applying the assumption that women without a positive OGTT did not have ODIP (67 women did have ODIP)
NPV = 0.999
PPV = 0.020
AUC = 0.86 (0.80 to 0.92)

The best RBG threshold was 7.31 -7.40mmol/l
Sensitivity = 0.78
Specificity = 0.85
LR = 5.2*
LR- = 0.26*

2) To estimate the maximum diagnostic value, using the assumption that those with no or incomplete OGTT and RBG < 11.1mmol/l did not have ODIP,</td>
<td>participants representative of the patients who will receive the test in practice: Yes
2) Were selection criteria clearly described: Yes
3) Was the reference standard likely to classify the target condition correctly: Yes
4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes
5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Only those with RBG test results >7.0mmol/l, or a previous history of gestational diabetes were tested using the reference standard in the first trimester. Those who were
normal according to the reference standard in the first trimester and
with GCT results > 7.7 mmol/l were tested using the reference standard in the second trimester
6) Did participants</td>
</tr>
</table>

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
2008 Source of funding Support from the National Institute for Health Research Cambridge Biomedical Research Centre	<table><tr><td>Head circumference - cm</td><td>34.7 (22.3-43.2) n = 11,483</td><td>34.8 (20.0-41.0) n = 5560</td></tr><tr><td>Ethnic origin: White British Asian African Caribbean Chinese Other White Backgrounds</td><td>12,725 (71.3) 703 (3.9) 133 (0.7) 63 (0.4) 205 (1.1) 3295 (18.5) n = 17,851</td><td>5465 (69.7) 349 (4.5) 92 (1.2) 35 (0.5) 112 (1.4) 1446 (18.4) n = 7841</td></tr><tr><td>Known maternal IV drug use</td><td>123 (1.0) n = 12,632</td><td>61 (1.0) n = 6211</td></tr><tr><td>Known maternal smoking in pregnancy</td><td>1654 (9.3) n = 17,845</td><td>777 (9.9) n = 7821</td></tr></table> Inclusion Criteria Women receiving antenatal and intrapartum care from East of England trust hospitals between 2004 and 2008 who had a live birth and for whom regional hospital obstetric data were available. Exclusion Criteria Women with recorded diabetes prior to pregnancy	glucose value ≥ 11.1 mmol/l.	hexokinase-glucose-6-phosphate dehydrogenase method.	but that 12 women who did not have an OGTT and had RBG ≥ 11.1mmol/l, did have ODIP NPV = 0.999 PPV = 0.028 AUC = 0.88 (0.83 to 0.93) The best RBG threshold was 7.51 - 7.59mmol/l Sensitivity = 0.80 Specificity = 0.88 LR = 6.67* LR- = 0.23* 3) To estimate the minimum diagnostic value, using only data from those women who had both RBG and OGTT performed (n=3007) (67 women had diagnosed ODIP) NPV = 0.988 PPV = 0.052 AUC = 0.72 (0.64 to 0.79) The best RBG threshold was 8.60 - 8.70 mmol/l Sensitivity = 0.60 Specificity = 0.75 LR = 2.4* LR- = 0.53* * Calculated by NCC-WCH	receive the same reference standard regardless of the index test result: No the 75g 2hour OGTT was always used as the reference standard but was not performed in all women who received the index test. Also venous or capillary samples were obtained but were analysed with reference only to the venous plasma glucose diagnostic criteri a values (capillary values are higher) 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: Yes 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
					standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Partially reported 14) Were withdrawals explained: Not relevant Other information
Full citation Corrado,F., D'anna,R., Cannata,M.L., Interdonato,M.L., Pintaudi,B., Di,Benedetto A., Correspondence between first-trimester fasting glycaemia, and oral glucose tolerance test in gestational diabetes diagnosis, Diabetes and	Sample size n=738/775 women (see exclusions below) Characteristics Characteristics of women are presented according to first trimester FPG result ≥ 5.1 mmol/l (n = 53) or < 5.1 mmol/l (n = 685) Age (years) FPG ≥ 5.1 mmol/l group = 30.63 ± 5.24 FPG < 5.1 mmol/l group = 33.42 ± 4.36 p = 0.0001 Prepregnancy BMI (kg/m2) FPG ≥ 5.1 mmol/l group = 23.8 ±7.32 FPG < 5.1 mmol/l group = 27.9 ± 5.81	Tests Screening test: FPG value from first trimester assay Diagnostic test: 2 hour 75g OGTT evaluated using IADPSG criteria (FPG >5.1mmol/l, 1 hour PG >10.0mmol/l, 2 hour PG >8.5mmol/l)	Methods All consecutive Caucasian women scheduled for an early third trimester 2 hour 75g OGTT were enrolled in the study. Pre-pregnancy BMI, age, parity and gestational age were noted. All women had been asked to provide the results of a first trimester	Results Incidence Overt DM using FPG (≥7mmol/l) in 1st tri = 6/744 (0.8%) Incidence of GDM using IADPSG/ADA 2011 75g OGTT in “early 3rd" trimester = 88/738 (12%) FPG Threshold at 5.1mmol/l in first trimester to detect gestational diabetes at week 24-28 TP: 24 FP: 29 FN: 64 TN: 621	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Metabolism, 38, 458-461, 2012 Ref Id 247650 Country/ies where the study was carried out Italy Study type Retrospective cohort study Aim of the study To evaluate the correspondence between first trimester fasting glycaemia and the results of the OGTT in diagnosing gestational diabetes using IADPSG criteria at 24-28 gestational weeks Study dates 2011 Source of funding Not stated	p = 0.0001 Gestational age (weeks) FPG ≥ 5.1 mmol/l group = 26.0 ± 2.7 FPG < 5.1 mmol/l group = 25.3 ±2.3 p = 0.064 Parity > 1 (n %) FPG ≥ 5.1 mmol/l group = 318/685 (46.4%) FPG < 5.1 mmol/l group = 31/53 (58.4%) p = 0.1 Prevalence of gestational diabetes (n %) FPG ≥ 5.1 mmol/l group = 64/685 (9.3%) FPG < 5.1 mmol/l group = 24/53 (45.3%) p = 0.0001 Inclusion Criteria Consecutive Caucasian pregnant women scheduled for an OGTT early in the third trimester of pregnancy Exclusion Criteria Twin pregnancy (n=12), no first trimester FPG assay (n=18), FPG value was determined after the first trimester (n=6), FPG diagnostic of pregestational diabetes >= 7 mmol/l (n=1)		FPG test (available free of charge). Women who had FPG<7mmol/l underwent an OGTT	Sensitivity,% (95% CI): 27.3 (19.7 - 35.0)* Specificity, % (95% CI): 95.5 (94.5 - 96.6)* LR+ (95% CI): 6.11 (3.59 - 10.25)* LR- (95% CI): 0.76 (0.67 - 0.85)* *Diagnostic accuracy measures and CIs calculated by NCC-WCH technical team based on data reported in the article	correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: Yes. Index test was a FPG test that was not performed as part of the 2hr 75g OGTT 8) Was the execution of the index test described in sufficient detail to permit its replication: Unclear 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Unclear

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
					10) Were index test results interpreted without knowledge of the results of the reference standard: Yes 11) Were the reference standard results interpreted without knowledge of the results of the index test: Yes 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: NA 14) Were withdrawals explained: NA Other information
Full citation Zhu,W.W., Yang,H.X., Wei,Y.M., Yan,J., Wang,Z.L., Li,X.L., Wu,H.R., Li,N., Zhang,M.H., Liu,X.H., Zhang,H., Wang,Y.H.,	Sample size n=17,186 medical records of pregnant women Characteristics Not stated Inclusion Criteria Pregnant women who received prenatal care at the GDM centers established in 13 hospitals in China	Tests Screening test: FPG test was performed at the first prenatal visit Diagnostic test: 75-g OGTT between 24 and 28 weeks' gestation	Methods In 13 hospitals in different parts of China, 17,186 pregnant women were tested for FPG at the first prenatal visit using venous blood sample collected after at least 8 h of	Results Incidence of gestational diabetes Incidence gestational diabetes using IADPSG 75g OGTT in 2nd trimester = 3002/17186 (17.4%) Diagnostic accuracy of FPG at 13.4 ± 3.5 weeks to detect gestational diabetes at 24-28 weeks using IADPSG criteria using 2 hour 75g OGTT	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Niu,J.M., Gan,Y.J., Zhong,L.R., Wang,Y.F., Kapur,A., Evaluation of the value of fasting plasma glucose in the first prenatal visit to diagnose gestational diabetes mellitus in china, Diabetes Care, 36, 586-590, 2013 Ref Id 247827 Country/ies where the study was carried out China Study type Retrospective cohort study Aim of the study To evaluate the value of fasting plasma glucose (FPG) value in the first prenatal visit to diagnose gestational diabetes mellitus Study dates	Exclusion Criteria Women with previously known diabetes were excluded from the study	evaluated using IADPSG criteria	fasting. Previously known diabetic patients were excluded from the study. For women with FPG ≥7.00 mmol/L at the first prenatal visit, medical care for diabetes was provided; for those with FPG <7.00 mmol/L, no interventions were made until women returned at 24 and 28 weeks in the fasting state for repeat testing, and this time a 75-g OGTT was performed. Venous blood samples were collected at 0, 1, and 2 h after a 75-g glucose load. iagnosis of gestational diabetes can be made when any one of the following values is met or exceeded in the 75-g OGTT: 0 h (fasting), ≥5.10 mmol/L; 1 h, ≥10.00 mmol/L;	FPG Threshold at 4.1mmol/l TP: 2816 FP: 12432 FN: 186 TN: 1752 Sensitivity,% (95% CI): 93.8 (92.9 - 94.6)* Specificity, % (95% CI): 12.4 (12.2 – 12.5)* LR+ (95% CI): 1.07 (1.06 – 1.08)* LR- (95% CI): 0.50 (0.43 – 0.58)* FPG Threshold at 4.6mmol/l TP: 1944 FP: 6259 FN: 1058 TN: 7935 Sensitivity,% (95% CI): 64.8 (63.2 – 66.3)* Specificity, % (95% CI): 55.9 (55.6 – 56.3)* LR+ (95% CI): 1.47 (1.42 – 1.52)* LR- (95% CI): 0.63 (0.60 – 0.66)* FPG Threshold at 5.1mmol/l TP: 779 FP: 1180 FN: 2223 TN: 13004 Sensitivity,% (95% CI): 25.9 (24.7 – 27.2)* Specificity, % (95% CI): 91.7 (91.4 – 92.0)* LR+ (95% CI): 3.12 (2.87 -3.38)* LR- (95% CI): 0.81 (0.79 – 0.82)* FPG Threshold at 5.6mmol/l TP: 162	2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Unclear 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: Yes. Index test was a FPG test that was not performed as part of the 2hr 75g OGTT 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
At Peking University First Hospital, records pertained to women registered at the prenatal clinic between 1 January 2010 and 31 December 2011 (the records after 1 May followed the new criteria), while at the other 12 participating hospitals, records pertained to women registered between 1 July 2011 and 29 February 2012. Source of funding World Diabetes Foundation			and 2 h, ≥8.50 mmol/L. Data of FPG at the first prenatal visit and 75-g OGTT at 24–28 weeks were analyzed.	FP: 129 FN: 2840 TN: 14055 Sensitivity,% (95% CI): 5.4 (4.8 – 5.9)* Specificity, % (95% CI): 99.1 (99.0 – 99.2)* LR+ (95% CI): 5.93 (4.7 - 7.5)* LR- (95% CI): 0.955 (0.95 -0.96)* FPG Threshold at 6.1 mmol/l TP: 43 FP: 12 FN: 2959 TN: 14172 Sensitivity,% (95% CI): 1.4 (1.2 – 1.6)* Specificity, % (95% CI): 99.9 (99.9 – 100)* LR+ (95% CI): 16.93 (8.65 – 33.83)* LR- (95% CI): 0.987 (0.98 – 0.99)* *Diagnostic accuracy measures and CIs calculated by NCC-WCH technical team based on data reported in the article	9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: No 14) Were withdrawals explained: Unclear Other information

A.13 Gestational diabetes - second trimester screening

<table>
<tr><th>Bibliographic details</th><th>Participants</th><th>Tests</th><th>Methods</th><th>Outcomes and results</th><th>Comments</th></tr>
<tr>
<td>Full citation
Agarwal,M.M., Dhatt,G.S., Punnose,J., Koster,G., Gestational diabetes in a high-risk population: using the fasting plasma glucose to simplify the diagnostic algorithm, European Journal of Obstetrics, Gynecology, and Reproductive Biology, 120, 39-44, 2005
Ref Id
179398
Country/ies where the study was carried out
United Arab Emirates
Study type
Prospective cohort study
Aim of the study</td>
<td>Sample size
1726 women attending routine antenatal clinics at Tawam Hospital

Characteristics
<table>
<tr><th>Characteristic</th><th>Women without gestational diabetes</th><th>Women with gestational diabetes</th><th>P value</th></tr>
<tr><td>n</td><td>1352 (80.2%)</td><td>333 (19.8%)</td><td></td></tr>
<tr><td>Age (years)</td><td></td><td></td><td></td></tr>
<tr><td>Mean ± SD</td><td>26.6 ± 5.7</td><td>29.3 ± 6.4</td><td>0.001</td></tr>
<tr><td>Median, Range</td><td>26, 16-48</td><td>28, 16-48</td><td></td></tr>
<tr><td>Gestational age at screening (weeks)</td><td></td><td></td><td></td></tr>
<tr><td>Mean ±SD</td><td>24.9 ± 5.3</td><td>25.2 ± 6.14</td><td>0.45</td></tr>
<tr><td>Median, Range</td><td>25, 9-40</td><td>25, 7-40</td><td></td></tr>
<tr><td>BMI</td><td></td><td></td><td></td></tr>
<tr><td>Mean ±SD</td><td>27.7 ± 8.5</td><td>28.9 ± 5.6</td><td>0.06</td></tr>
</table>
</td>
<td>Tests
Index test: FPG
Reference standard: 2 hour 75 gram oral glucose tolerance test
Diagnostic criteria: WHO 1999 thresholds for gestational diabetes - fasting plasma glucose value (FPG) ≥ 7 mmol/l and/or 2h postload plasma glucose value ≥ 7.8 mmol/l</td>
<td>Methods
For OGTT: Venous blood samples were collected for fasting and 1 and 2 hour post 75g oral glucose load after women had fasted overnight for 12 hours. Plasma glucose was estimated using the glucose oxidase method. The overall coefficient of variation was 3.7% and the hospital laboratory met standards for internal and external quality assurance.</td>
<td>Results
Incidence of gestational diabetes
Incidence of gestational diabetes in study population = 333/1685 (19.8%)

Diagnostic test accuracy of FPG index test at different thresholds compared with reference standard 2 hour OGTT interpreted using WHO 1999 criteria thresholds (FPG ≥ 7.0 or 2 hour PG ≥ 7.8 mmol/l)
at FPG test threshold of 3.9mmol/l
TP: 332* FN: 1* FP: 1348* TN: 4*
Sensitivity, % (95% CI): 99.7 (98.9 to 100)*
Specificity, % (95% CI): 0.3 (0.1 to 0.4)*
LR (95% CI): 1.00 (0.99 to 1.00)*
LR- (95% CI): 1.02 (0.04 to 9.50)*

at FPG test threshold of 4.2 mmol/l
TP: 325* FN: 8* FP: 1308* TN: 44*
Sensitivity, % (95% CI): 97.6 (95.6 to 98.8)*
Specificity, % (95% CI): 3.3 (2.8 to 3.6)*
LR (95% CI): 1.01 (0.98 to</td>
<td>Limitations
NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy
1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes
2) Were selection criteria clearly described: Yes
3) Was the reference standard likely to classify the target condition correctly: Yes
4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes
5) Did the whole sample or a random selection of the sample receive verification using the reference standard:</td>
</tr>
</table>

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
To evaluate fasting plasma glucose (FPG) as a screening test for gestational diabetes Study dates 1 June 2003 to 31 January 2004 Source of funding None stated	Inclusion Criteria Women attending routine antenatal clinics at Tawam Hospital, Al Ain who received universal screening Exclusion Criteria Women who were unable to complete the oral glucose tolerance test (OGTT) due to vomiting, refusal to undergo testing, who ate or drank during the test or other reasons (n = 41)			1.03)* LR- (95% CI): 0.74 (0.32 to 1.61)* at FPG test threshold of 4.4 mmol/l TP: 311* FN: 22* FP: 1196* TN: 156* Sensitivity, % (95% CI): 93.4 (90.4 to 95.6)* Specificity, % (95% CI): 11.5 (10.8 to 12.1)* LR (95% CI): 1.06 (1.01 to 1.09)* LR- (95% CI): 0.57 (0.36 to 0.89)* at FPG test threshold of 4.7 mmol/l TP: 260* FN: 917* FP: 73* TN: 435* Sensitivity, % (95% CI): 78.1 (73.6 to 82.0)* Specificity, % (95% CI): 32.2 (31.1 to 33.2)* LR (95% CI): 1.15 (1.07 to 1.23)* LR- (95% CI): 0.68 (0.54 to 0.85)* at FPG test threshold of 5 mmol/l TP: 194* FN: 139* FP: 499* TN: 853* Sensitivity, % (95% CI): 58.3 (53.3 to 63.0)* Specificity, % (95% CI): 63.1 (61.9 to 64.3)* LR (95% CI): 1.58 (1.34 to	The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No, the index test did form part of the reference standard 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				1.76)* LR- (95% CI): 0.66 (0.58 to 0.75)* at FPG test threshold of 5.3 mmol/l TP: 125* FN: 208* FP: 223* TN: 1129* Sensitivity, % (95% CI): 37.5 (33.1 to 42.1)* Specificity, % (95% CI): 83.5 (82.4 to 84.6)* LR (95% CI): 2.28 (1.88 to 2.74)* LR- (95% CI): 0.75 (0.69 to 0.81)* at FPG test threshold of 5.6 mmol/l TP: 80* FN: 253* FP: 93* TN: 1259* Sensitivity, % (95% CI): 24.0 (20.4 to 27.7)* Specificity, % (95% CI): 93.1 (92.2 to 94.0)* LR (95% CI): 3.49 (2.63 to 4.63)* LR- (95% CI): 0.82 (0.77 to 0.86)* at FPG test threshold of 5.8 mmol/l TP: 58* FN: 275* FP: 44* TN: 1308* Sensitivity, % (95% CI): 17.4 (14.4 to 20.2)* Specificity, % (95% CI): 96.7 (96.0 to 97.4)* LR (95% CI): 5.35 (3.63 to	when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: No 14) Were withdrawals explained: Yes Other information

<table>
<tr><th>Bibliographic details</th><th>Participants</th><th>Tests</th><th>Methods</th><th>Outcomes and results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td>7.92)*
LR- (95% CI): 0.85 (0.82 to 0.89)*

at FPG test threshold of 6.1 mmol/l
TP: 30* FN: 303* FP: 11* TN: 1341*
Sensitivity, % (95% CI): 9.0 (7.0 to 10.5)*
Specificity, % (95% CI): 99.2 (98.7 to 99.5)*
LR (95% CI): 11.07 (5.40 to 23.3)*
LR- (95% CI): 0.92 (0.90 to 0.94)*

TP - true positive, FN - false negative, FP - false positive, TN - true negative
* Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html</td><td></td></tr>
<tr><td>Full citation
Agarwal,M.M., Dhatt,G.S., Punnose,J., Koster,G., Gestational diabetes: a reappraisal of HBA_{1c} as a screening test, Acta Obstetricia et Gynecologica Scandinavica,</td><td>Sample size
454 women attending routine antenatal clinical at Tawam Hospital Al Ain and receiving universal screening

Characteristics<table><tr><th>Characteristic</th><th>Women without gestational diabetes</th><th>Women with gestational diabetes</th><th>p value</th></tr><tr><td>n</td><td>358</td><td>84</td><td></td></tr></table></td><td>Tests
Index test: HbA_{1c}
Reference standard: 2 hour 75 gram oral glucose tolerance test
Diagnostic criteria: WHO 1999 thresholds for gestational diabetes - fasting plasma glucose value (FPG) ≥ 7 mmol/l and/or 2h postload plasma glucose value ≥ 7.8 mmol/l</td><td>Methods
For HbA_{1c}: An EDTA sample for HbA_{1c} was collected together with the fasting glucose sample and was measured using an automated turbidimeteric immunoinhibition method. The coefficient of</td><td>Results
Incidence of gestational diabetes
Incidence of gestational diabetes in study population = 84/442 (19%)

Diagnostic test accuracy of HbA_{1c} index test at different thresholds compared with reference standard 2 hour OGTT interpreted using WHO</td><td>Limitations
NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy
1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes
2) Were selection</td></tr>
</table>

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
84, 1159-1163, 2005 Ref Id 179397 Country/ies where the study was carried out United Arab Emirates Study type Prospective cohort study Aim of the study To evaluate HbA_{1c} as a screening test for gestational diabetes Study dates 1 May to 31 July 2003 Source of funding None stated	<table><tr><td>Age (years)</td><td></td><td></td><td></td></tr><tr><td>Mean ± SD</td><td>26.15 ± 5.3</td><td>28.5 ± 5.9</td><td>0.001</td></tr><tr><td>Median, Range</td><td>25, 16-48</td><td>27.5, 16-42</td><td></td></tr><tr><td>Gestational age at screening (weeks)</td><td></td><td></td><td></td></tr><tr><td>Mean ± SD</td><td>26 ± 4.5</td><td>27 ± 4.85</td><td>0.003</td></tr><tr><td>Median, Range</td><td>25, 16-40</td><td>28, 18-37</td><td></td></tr><tr><td>Ethnic Group (%)</td><td></td><td></td><td></td></tr><tr><td>UAE arabs</td><td>244 (68.2)</td><td>57 (67.9)</td><td>0.7</td></tr><tr><td>Asian arabs</td><td>62 (17.3)</td><td>16 (19)</td><td></td></tr><tr><td>Chami arabs</td><td>12 (3.4)</td><td>1 (1.2)</td><td></td></tr><tr><td>East African arabs</td><td>4(1.1)</td><td>1 (1.2)</td><td></td></tr><tr><td>Indian subcontinent</td><td>5 (1.4)</td><td>2 (2.4)</td><td></td></tr><tr><td>Other</td><td>7 (1.9)</td><td>0 (0)</td><td></td></tr><tr><td>Unknown</td><td>24 (6.7)</td><td>7 (8.3)</td><td></td></tr></table> Inclusion Criteria Women attending routine antenatal clinical at Tawam Hospital Al Ain who received universal screening Exclusion Criteria Women who were unable to complete the OGTT due to vomiting (n = 12)		variation was 3.2% and the hospital laboratory met standards for internal and external quality assurance. For OGTT: Venous blood samples were collected for fasting and 1 and 2 hour post 75g oral glucose load after women had fasted overnight for 12 hours. Plasma glucose was estimated using the glucose oxidase method. The overall coefficient of variation was 2% and the hospital laboratory met standards for internal and external quality assurance.	1999 criteria thresholds (FPG ≥ 7.0 or 2 hour PG ≥ 7.8 mmol/l) at HbA_{1c} test threshold of 4.5% TP: 82* FN: 2* FP: 353* TN: 5* Sensitivity, % (95% CI): 97.6 (94.2 to 99.6)* Specificity, % (95% CI): 1.4 (0.6 to 1.9)* LR (95% CI): 0.99 (0.95 to 1.02)* LR- (95% CI): 1.70 (0.23 to 9.69)* at HbA_{1c} test threshold of 5% TP: 82* FN: 2* FP: 341* TN:17* Sensitivity, % (95% CI): 97.6 (94.2 to 99.6)* Specificity, % (95% CI): 4.7 (3.5 to 5.2)* LR (95% CI): 1.02 (0.96 to 1.05)* LR- (95% CI): 0.50 (0.08 to 2.17)* at HbA_{1c} test threshold of 5.5% TP: 69* FN: 15* FP: 283* TN: 75* Sensitivity, % (95% CI): 82.1 (73.2 to 89.0)* Specificity, % (95% CI): 20.9 (18.9 to 22.6)* LR (95% CI): 1.04 (0.90 to 1.15)*	criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: Yes 8) Was the execution of the index test described in sufficient

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				LR- (95% CI): 0.85 (0.49 to 1.42)* at HbA_{1c} test threshold of 6% TP: 41* FN: 43* FP: 159* TN: 199* Sensitivity, % (95% CI): 48.8 (38.8 to 58.9)* Specificity, % (95% CI): 55.6 (53.2 to 57.9)* LR (95% CI): 1.10 (0.83 to 1.40)* LR- (95% CI): 0.92 (0.71 to 1.15)* at HbA_{1c} test threshold of 6.5% TP: 18* FN: 66* FP: 77* TN: 281* Sensitivity, % (95% CI): 21.4 (13.9 to 30.6)* Specificity, % (95% CI): 78.5 (76.7 to 80.6)* LR (95% CI): 1.00 (0.60 to 1.58)* LR- (95% CI): 1.00 (0.86 to 1.12)* at HbA_{1c} test threshold of 7% TP: 9* FN: 75* FP: 34* TN: 324* Sensitivity, % (95% CI): 10.7 (5.5 to 18.1)* Specificity, % (95% CI): 90.5 (89.3 to 92.2)* LR (95% CI): 1.13 (0.52 to 2.32)*	detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: No 14) Were withdrawals explained: Yes Other information

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				LR- (95% CI): 0.99 (0.89 to 1.06)* at HbA_{1c} test threshold of 7.5% TP: 6* FN: 78* FP: 15* TN: 343* Sensitivity, % (95% CI): 7.1 (3.1 to 12.9)* Specificity, % (95% CI): 95.8 (94.9 to 97.2)* LR (95% CI): 1.70 (0.60 to 4.51)* LR- (95% CI): 0.97 (0.90 to 1.02)* at HbA_{1c} test threshold of 8% TP: 3* FN: 81* FP: 5* TN: 353* Sensitivity, % (95% CI): 3.6 (1.0 to 7.0)* Specificity, % (95% CI): 98.6 (98.0 to 99.4)* LR (95% CI): 2.56 (0.49 to 12.03)* LR- (95% CI): 0.98 (0.94 to 1.01)* TP - true positive, FN - false negative, FP - false positive, TN - true negative * Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html	
Full citation	Sample size	Tests	Methods	Results	Limitations

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Agarwal,M.M., Dhatt,G.S., Punnose,J., Gestational diabetes: utility of fasting plasma glucose as a screening test depends on the diagnostic criteria, Diabetic Medicine, 23, 1319-1326, 2006 Ref Id 152942 Country/ies where the study was carried out United Arab Emirates Study type Prospective cohort study Aim of the study To estimate the effect of diagnostic criteria on the performance of fasting plasma glucose (FPG) as a	4844 women attending routine antenatal clinic at Al Ain Hospital Characteristics Mean maternal age = 28.4 years (median 28 years SD 6.0, range 16-48 years) Ethnicity: 3473 (75.5%) Arab, 932 (20.3%) South Asian (India, Pakistan, Bangladesh and Sri Lanka), 92 (2%) Other nationalities, 105 (2.3%) unavailable Mean gestational age at oral glucose tolerance test (OGTT) = 25.9 gestational weeks (median 26 weeks, SD 6.3, range 2-38 weeks) Inclusion Criteria All women attending routine antenatal clinic at Al Ain Hospital who underwent a 75g OGTT as part of a universal screening programme Exclusion Criteria 242 women who did not undergo 75g OGTT because of refusal (n = 242), vomiting during the test (n = 110) or eating food during the test of other reasons (n = 17). A further 74 women who were diagnosed with gestational diabetes on the basis of FPG results alone were excluded from the published analyses, but were included in the analysis for this review	Index test: Fasting plasma glucose Reference standard: 75g OGTT Diagnostic criteria: WHO 1999 thresholds for gestational diabetes - FPG ≥ 7mmol/l and/or 2 h postload glucose value ≥ 7.8 mmol/l	For OGTT: Following a 12 hour overnight fast, venous plasma samples were collected fasting and 1 and 2 hours after an oral 75g glucose load. Plasma glucose was determined using the glucose oxidase method. The overall coefficient of variation was 2.4% and the hospital laboratory met standards for internal and external quality assurance for glucose measurment	Incidence of gestational diabetes Incidence of gestational diabetes in second trimester at gestational week 24-28 = 979/4596 (21.3%)* Diagnostic test accuracy of FPG index test at different thresholds compared with reference standard 2 hour OGTT interpreted using WHO 1999 criteria thresholds (FPG ≥ 7.0 or 2 hour PG ≥ 7.8 mmol/l) at FPG test threshold of 4.2 mmol/l TP: 930* FN: 55* FP: 3242* TN: 375* Sensitivity, % (95% CI): 94.4 (92.9 to 95.7)* Specificity, % (95% CI): 10.4 (10.0 to 10.7)* LR (95% CI): 1.05 (1.03 to 1.07)* LR- (95% CI): 0.54 (0.40 to 0.71)* at FPG test threshold of 4.4 mmol/l TP: 856* FN: 128* FP: 2575* TN: 1043* Sensitivity, % (95% CI): 87.0 (84.9 to 88.9)* Specificity, % (95% CI): 28.8 (28.3 to 29.3)* LR (95% CI): 1.22 (1.18 to 1.25)*	NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
screening test for gestational diabetes Study dates May 2004 to September 2005 Source of funding None stated				LR- (95% CI): 0.45 (0.38 to 0.54)* at FPG test threshold of 4.7 mmol/l TP: 706* FN: 279* FP: 1752* TN: 1865* Sensitivity, % (95% CI): 71.7 (69.0 to 74.2)* Specificity, % (95% CI): 51.6 (50.8 to 52.3)* LR (95% CI): 1.48 (1.40 to 1.55)* LR- (95% CI): 0.55 (0.49 to 0.61)* at FPG test threshold of 5.0 mmol/l TP: 545* FN: 439* FP: 965* TN: 2653* Sensitivity, % (95% CI): 55.4 (52.6 to 58.1)* Specificity, % (95% CI): 73.3 (72.6 to 74.1)* LR (95% CI): 2.08 (1.92 to 2.24)* LR- (95% CI): 0.61 (0.57 to 0.65)* at FPG test threshold of 5.3 mmol/l TP: 402* FN: 583* FP: 485* TN: 3132* Sensitivity, % (95% CI): 40.8 (38.3 to 43.3)* Specificity, % (95% CI): 86.6 (85.9 to 87.3)* LR (95% CI): 3.04 (2.72 to 3.40)*	reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No, the index test did form part of the reference standard 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				LR- (95% CI): 0.68 (0.65 to 0.72)* at FPG test threshold of 5.6 mmol/l TP: 293* FN: 691* FP: 206* TN: 3412* Sensitivity, % (95% CI): 29.8 (27.7 to 31.8)* Specificity, % (95% CI): 94.3 (93.7 to 94.9)* LR (95% CI): 5.23 (4.43 to 6.18)* LR- (95% CI): 0.74 (0.72 to 0.77)* at FPG test threshold of 5.8 mmol/l TP: 218* FN: 768* FP: 93* TN: 3523* Sensitivity, % (95% CI): 22.1 (20.5 to 23.6)* Specificity, % (95% CI): 97.4 (97.0 to 97.8)* LR (95% CI): 8.60 (6.78 to 10.92)* LR- (95% CI): 0.80 (0.78 to 0.82)* TP - True positive, FN - false negative, FP - false positive, TN - true negative * Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html	when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: No 14) Were withdrawals explained: Yes Other information

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation Agarwal,M.M., Dhatt,G.S., Shah,S.M., Gestational diabetes mellitus: simplifying the international association of diabetes and pregnancy diagnostic algorithm using fasting plasma glucose, Diabetes Care, 33, 2018-2020, 2010 Ref Id 153971 Country/ies where the study was carried out United Arab Emirates Study type Retrospective cohort study Aim of the study To determine the effect of the	Sample size Data from 10,283 women were available for analysis Characteristics The baseline characteristics of participants are not described in detail. Ethnicity: 8233 (80.1%) were of Arab ethnicity and 1592 (15.5%) were of South Asian ethnicity Inclusion Criteria Participants from four previous studies by the authors were included. These women attended routine antenatal clinics at two tertiary care hospitals and underwent a 75g oral glucose tolerance test (OGTT) at gestational weeks 24-28 as part of a universal screening programme. No further details are provided. Exclusion Criteria No details are provided	Tests Index test: Fasting plasma glucose (FPG) Reference standard: 75g OGTT performed at gestational weeks 24-28 Diagnostic criteria: IADPSG thresholds for gestational diabetes - one or more plasma venous glucose values FPG ≥ 5.1mmol/l, 1 hour ≥ 10.0mmol/l or 2 hour ≥ 8.5mmol/l	Methods For OGTT: Plasma glucose was estimated using the glucose oxidase method and analytical standards for glucose were met.	Results Incidence of gestational diabetes Incidence at 24-28 weeks = 3875/10283 (37.7%) Diagnostic test accuracy of FPG index test at different thresholds compared with reference standard 2 hour OGTT interpreted using IADPSG criteria thresholds (FPG ≥ 5.1 and/or 1 hour PG ≥ 10.0 mmol/l and/or 2 hour PG ≥ 8.5 mmol/l) at FPG test threshold of 4.2 mmol/l TP: 3809* FN: 66* FP: 5669* TN: 739* Sensitivity, % (95% CI): 98.3 (97.9 to 98.7)* Specificity, % (95% CI): 11.5 (11.3 to 11.8)* LR (95% CI): 1.11 (1.10 to 1.12)* LR- (95% CI): 0.15 (0.11 to 0.19)* at FPG test threshold of 4.4 mmol/l TP: 3697* FN: 178* FP: 4358* TN: 2050* Sensitivity, % (95% CI): 95.4 (94.7 to 96.0)* Specificity, % (95% CI): 32.0 (31.6 to 32.4)* LR (95% CI): 1.40 (1.38 to 1.42)* LR- (95% CI): 0.14 (0.12 to	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No, exclusion criteria not described 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard:

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
International Association of Diabetes and Pregnancy Study Group (IADPSG) criteria on gestational diabetes diagnosis and the fasting plasma glucose to predict gestational diabetes Study dates Data from four studies conducted between 2003 to 2008 were reanalysed using IADPSG criteria Source of funding None stated				0.17)* at FPG test threshold of 4.7 mmol/l TP: 3445* FN: 430* FP: 2555* TN: 3853* Sensitivity, % (95% CI): 88.9 (88.0 to 89.8)* Specificity, % (95% CI): 60.1 (59.6 to 60.7)* LR (95% CI): 2.23 (2.18 to 2.28)* LR- (95% CI): 0.19 (0.17 to 0.20)* at FPG test threshold of 5.0 mmol/l TP: 3119* FN: 756* FP: 582* TN: 5826* Sensitivity, % (95% CI): 80.5 (79.6 to 81.3)* Specificity, % (95% CI): 90.9 (90.4 to 91.4)* LR (95% CI): 8.86 (8.28 to 9.49)* LR- (95% CI): 0.22 (0.20 to 0.23)* at FPG test threshold of 5.1 mmol/l TP: 2975* FN: 900* FP: 0* TN: 6408* Sensitivity, % (95% CI): 76.77 (75.42 to 78.08)** Specificity, % (95% CI): 99.99 (99.94 to 100)** LR (95% CI): 9840 (872 to 5159878830)** LR- (95% CI): 0.232 (0.232	The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No, the index test did form part of the reference standard 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available

<table>
<tr><th>Bibliographic details</th><th>Participants</th><th>Tests</th><th>Methods</th><th>Outcomes and results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td>to 0.234)**

TP - true positive, FN - false negative, FP - false positive, TN - true negative
* Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html

** 0.5 has been added to each cell (TP, FN, FP, TN) for diagnostic accuracy calculations to take into account the zeros</td><td>when the test results were interpreted as would be available when the test is used in practice: Yes
13) Were uninterpretable, indeterminate or intermediate test results reported: No
14) Were withdrawals explained: Yes

Other information</td></tr>
<tr><td>Full citation
Bito,T., Nyari,T., Kovacs,L., Pal,A., Oral glucose tolerance testing at gestational weeks < or =16 could predict or exclude subsequent gestational diabetes mellitus during the current pregnancy in high risk group, European</td><td>Sample size
163 women at 16 gestational weeks or less were enrolled in the study. Women with gestational diabetes diagnosed at 16 gestational weeks or less were excluded from the study (n = 8)

Characteristics<table><tr><th></th><th>Onset of gestational diabetes at weeks 24-28</th><th>No. gestational diabetes at weeks 24-28 and weeks 32-34</th><th>Total (includes women wi gestational diabetes a weeks 32-34)</th></tr><tr><td>Mean age (years)</td><td>30.2 ± 4.9</td><td>28.1 ± 5.3</td><td>28.7 ± 5.2</td></tr><tr><td>Mean BMI (kg/m2)</td><td>28.4 ± 7.3</td><td>25.3.1 ± 4.4</td><td>26.7 ± 5.6</td></tr><tr><td>Mean glucose</td><td>5.4 ± 0.7</td><td>4.6 ± 0.4</td><td>4.9 ± 0.6</td></tr></table></td><td>Tests
Index test: No index test was used
Reference standard: 2 hour 75 gram OGTT
Diagnostic criteria: WHO 1999 thresholds for gestational diabetes - fasting plasma glucose value (FPG) ≥ 7 mmol/l and/or 2h postload plasma glucose value (2h PG) ≥ 7.8 mmol/l</td><td>Methods
For OGTT: Women were instructed to consume at least 150g of carbohydrate each day for 3 days and then to adhere to a 10-12 hour overnight fast the day before the OGTT. Venous plasma samples were collected at fasting and 2 hours after ingestion of 75g glucose solution over a 5 minute period. Glucose levels were determined by the</td><td>Results
Incidence of gestational diabetes
Incidence of gestational diabetes in second trimester
at gestational week 24-28 = 32/155 (20.64%)*

Incidence of gestational diabetes in second trimester/ Incidence of gestational diabetes
by gestational week 24-28 = 32/40 (80%)*

Diagnostic test accuracy of FPG index test at threshold of 5.0 mmol/l compared with reference standard 2 hour OGTT interpreted using WHO 1999 criteria thresholds</td><td>Limitations
NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy
1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes
2) Were selection criteria clearly described: Yes
3) Was the reference standard likely to classify the target condition correctly: Yes
4) Was the period between performance</td></tr>
</table>

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Journal of Obstetrics, Gynecology, and Reproductive Biology, 121, 51-55, 2005 Ref Id 152996 Country/ies where the study was carried out Hungary Study type Prospective cohort study Aim of the study To determine possible upper and lower cut-off values for the oral glucose tolerance test (OGTT) at or before gestational week 16 to predict subsequent onset of gestational diabetes in a high risk population, to assess the proportion of	<table><tr><td>level at fasting</td><td></td><td></td><td></td></tr><tr><td>Mean glucose level at 120 mins postload</td><td>7.1 ± 0.4</td><td>5.5 ± 1.0</td><td>6.1 ± 1.1</td></tr><tr><td>No. (%) cases with 1 risk factor</td><td>19 (59.4%)</td><td>60 (80%)</td><td>109 (70.3%</td></tr><tr><td>No. (%) cases with ≥ 2 risk factors</td><td>13 (40.6%)</td><td>15 (20%)</td><td>46 (29.7%)</td></tr></table> Inclusion Criteria All pregnant women who did not have a previous history of gestational diabetes or any history of alteration of carbohydrate metabolism, but who displayed one or more risk factors for gestational diabetes and who were referred to the specialist outpatient department. The risk factors were: family history of type 2 diabetes, history of a large neonate (≥ 4000g), history of an adverse perinatal outcomes (missed abortion, malformation, polyhydramnios, stillbirth or preterm delivery), obesity (pre-pregnant BMI ≥ 30m2), age ≥ 35 years or glycosuria. Exclusion Criteria Women who were diagnosed as having gestational diabetes by OGTT at < 16 gestational weeks were excluded from the study		GOD-POD colorimetric method on sodium fluoride-mediated blood. The interassay and the interassay coefficient of variation were < 2%.	(FPG ≥ 7.0 or 2 hour PG ≥ 7.8 mmol/l) TP: 29* FN: 3* FP: 88* TN: 35* Sensitivity, % (95% CI): 90.6 (75.8 to 97.5)* Specificity, % (95% CI): 28.5 (24.6 to 30.2)* LR (95% CI): 1.27 (1.01 to 1.40)* LR- (95% CI): 0.33 (0.08 to 0.98)* TP - true positive, FN - false negative, FP - false positive, TN - true negative * Calculated by NCC-WCH	of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No, the index test did form part of the reference standard 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication:

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
the group that would not require further OGTTs if these were applied and to determine the predictive values for different risk factors for gestational diabetes at gestational weeks 24-28 and 32-34. Study dates 1 January 2001 to 30 September 2002 Source of funding Not stated					Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: No 14) Were withdrawals explained: There were no withdrawals Other information
Full citation Black,M.H., Sacks,D.A., Xiang,A.H., Lawrence,J.M., Clinical outcomes of pregnancies	Sample size 9199 women atending the KPSC Bellflower Medical Centre Characteristics Maternal Characteristic All women No gestationa	Tests Index test: none Reference standard: 75g 2 hour OGTT Diagnostic criteria: IADPSG thresholds for gestational diabetes - one or	Methods No details are provided regarding the laboratory methods and standards of glucose testing.	Results Incidence of gestational diabetes Incidence of gestational diabetes in whole study population = 2179/9199 (23.7%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum

Bibliographic details	Participants				Tests	Methods	Outcomes and results	Comments
complicated by mild gestational diabetes mellitus differ by combinations of abnormal oral glucose tolerance test values, Diabetes Care, 33, 2524-2530, 2010 Ref Id 178358 Country/ies where the study was carried out USA Study type Retrospective cohort study Aim of the study To examine the association between the different glucose values assessed within the oral glucose tolerance test (fasting, 1	n	8711	7020	1691	more plasma venous glucose values FPG ≥ 5.1mmol/l, 1 hour ≥10.0mmol/l or 2 hour ≥ 8.5mmol/l		Incidence of gestational diabetes in untreated study population = 1691/8711(19.4%) Incidence of adverse outcomes Large for gestational age (Definition: infants in whom sex-specific,race-specific and gestational age-specific birth wieght > 90th percentile) No gestational diabetes = 528/7020 Gestational diabetes = 264/1691 RR (95% CI) = 2.08 (1.80 to 2.38) P < 0.0001 Primary ceasarean section (Confirmed from infant birth certificate) No gestational diabetes = 1112/7020 Gestational diabetes = 336/1691 RR (95% CI) = 0.96 (0.87 to 1.07) P = 0.49 Shoulder dystocia/birth injury (Definition: ICD-9 codes 653.4, 653.5, 660.4, 767.0 - 767.9 or 959.0 - 959.9 at delivery) No gestational diabetes = 268/7020 Gestational diabetes =	of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Reference standard used only 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: reference standard used only, no index test used
	Race/ethnicity (%)							
	Non-Hispanic white	626 (7.2)	507 (7.2)	119 (7.0)				
	Hispanic	6484 (74.4)	5216 (74.3)	1268 (75.0)				
	Black	880 (10.1)	741 (10.6)	139 (8.2)				
	Asian	641 (6.4)	493 (7.0)	148 (8.8)				
	Other	80 (0.9)	63 (0.9)	17 (1.0)				
	Parity (%)							
	0	3492 (40.1)	2924 (41.7)	568 (33.6)				
	1	2675 (30.7)	2151 (30.6)	524 (31.0)				
	≥ 2	2479 (28.5)	1888 (26.9)	591 (35.0)				
	Unknown	65 (0.7)	57 (0.8)	8 (0.5)				
	Pregravid BMI (kg/m2)							
	Normal	3497 (40.1)	3096 (44.1)	401 (23.7)				
	Overweight	2733 (31.4)	2187 (31.2)	546 (32.3)				
	Obese	2481 (28.5)	1737 (24.7)	744 (44.0)				
	Prenatal smoking (%)							

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
hour and 2 hour plasma values) and adverse maternal and perinatal outcomes in untreated women, accounting for differences in maternal demographics, pre-pregnancy BMI and gestational weight gain. Also, to investigate associations between adverse outcomes and different categories of hyperglycaemia that result in a diagnosis of gestational diabetes using International Association of Diabetes in Pregnancy Study Groups (IADPSG) criteria to	<table><tr><td>No</td><td>8031 (92.2)</td><td>6490 (92.4)</td><td>1542 (91.1)</td></tr><tr><td>Yes</td><td>217 (2.5)</td><td>172 (2.5)</td><td>25 (2.7)</td></tr><tr><td>Unknown</td><td>463 (5.3)</td><td>358 (5.1)</td><td>105 (6.2)</td></tr><tr><td>Infant Characteristic</td><td></td><td></td><td></td></tr><tr><td>Preterm delivery</td><td>638* (7.3)</td><td>465 (6.6)</td><td>173 (10.2)</td></tr></table> * Calculated by NCC-WCH Inclusion Criteria Women who had a live singeton birth at ≥ 20 weeks gestation at the KPSC Bellflower Medical Centre within the study period, who had a prenatal 2 hour 75g OGTT with no prior 50g oral glucose challenge test, for whom pre-pregnancy and delivery anthropometric data were available and who did not receive treatment Exclusion Criteria Women receiving any form of treatment during pregnancy (n = 488). Only data from the first birth were included for women who had more than one birth during the study period.			96/1691 RR (95% CI) = 1.09 (0.88 to 1.36) P = 0.42	7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No index test used 8) Was the execution of the index test described in sufficient detail to permit its replication: No index test used 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: No index test used 11) Were the reference standard results interpreted without knowledge of the results of the index test: No index test used 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were

<table>
<tr><th>Bibliographic details</th><th>Participants</th><th>Tests</th><th>Methods</th><th>Outcomes and results</th><th>Comments</th></tr>
<tr><td>assess whether the level of risk is similar for individual and combinations of oral glucose tolerance test (OGTT) results.

Study dates
1 October 2005 to 31 March 2010

Source of funding
Supported by Kaiser Permanente Southern California Direct Community Benefit Funds</td><td></td><td></td><td></td><td></td><td>uninterpretable, indeterminate or intermediate test results reported: No
14) Were withdrawals explained: Yes

Other information</td></tr>
<tr><td>Full citation
Catalano,P.M., McIntyre,H.D., Cruickshank, J.K., McCance,D.R., Dyer,A.R., Metzger,B.E., Lowe,L.P., Trimble,E.R.,</td><td>Sample size
53,295 women from 15 centres in nine countries were eligible to participate. 28,562 (53.6%) agreed to take part in the study and 25,505 women completed the oral glucose tolerance test (OGTT). Data from 23,316 women were available for analysis.

Characteristics<table><tr><th>Characteristics</th><th>N</th><th>%</th><th>Mean</th><th>SD</th></tr><tr><td>Maternal</td><td></td><td></td><td></td><td></td></tr></table></td><td>Tests
Index test: none
Reference standard: 75g 2 hour OGTT
Diagnostic criteria: International Association of Diabetes and Pregnancy Study Group (IADPSG) thresholds for</td><td>Methods
To examine the associations of gestational diabetes and obesity, singly and in combination, HAPO participants were divided into four mutually exclusive groups: 1) no gestational</td><td>Results
Incidence of gestational diabetes
Incidence of gestational diabetes in study population
= 3746/23267* (16.1%)

Incidence of adverse outcomes
Birthweight > 90th</td><td>Limitations
NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy
1) Was the spectrum of participants representative of the patients who will</td></tr>
</table>

Bibliographic details	Participants					Tests	Methods	Outcomes and results	Comments
Coustan,D.R., Hadden,D.R., Persson,B., Hod,M., Oats,J.J.N., The hyperglycemia and adverse pregnancy outcome study: Associations of GDM and obesity with pregnancy outcomes, Diabetes Care, 35, 780-786, 2012 Ref Id 181728 Country/ies where the study was carried out International study: USA, Australia, UK and Isreal Study type Prospective cohort study Aim of the study To examine associations	Age (years) 23,316 29.2 5.8 BMI (kg/m2) 23,316 27.7 5.1 Gestational age (weeks) 23,316 27.8 1.8 Pre pregnant BMI 21,324 23.9 5.0 Ethnicity White, non-Hispanic 11,265 48.3 Black, non-Hispanic 2,696 11.6 Hispanic 1,984 8.5 Asian 6,757 29.0 Other 614 2.6 Parity (prior delivery ≥20 weeks) 12,233 52.5 Any prenatal smoking 1,581 6.8 Family history of diabetes 5,282 22.7 Obese 3,198 13.7 Overweight 5,143 22.1 Normal weight, underweight 14,975 64.2 Inclusion Criteria All pregnant women at each field centre were eligible to participate unless they had one or more exclusion criteria (not published here but published previously) Exclusion Criteria					gestational diabetes - one or more plasma venous glucose values FPG ≥ 5.1mmol/l, 1 hour ≥ 10.0mmol/l or 2 hour ≥ 8.5mmol/l	diabetes, no obesity; 2) gestational diabetes, no obesity; 3) no gestational diabetes, obesity; and 4) gestational diabetes, obesity. Two logistic regression models were then fit for each outcome (not presented here), with no gestational diabetes and no obesity used as the referent group. No details are presented regarding performance of the OGTT	percentile (Definition: The 90th percentile was considered to be present if the birth weight was greater than the 90th percentile for the baby's sex, gestational age, ethnicity, field centre, and maternal parity with gestational ages of 30–44 weeks included) Entire population No gestational diabetes = 1617/19491 (8.3%) Gestational diabetes = 604/3726 (16.2%) RR (95% CI) = RR 1.95 (1.79 to 2.13) P < 0.00001 Obese women No gestational diabetes = 278/2247 (12.4%) Gestational diabetes = 203/935 (21.7%) RR (95% CI) = RR 1.75 (1.49 to 2.07) P < 0.00001 Cord C-peptide > 90th percentile (Definition: Cord C-peptide > 90th percentile Cord blood was collected at delivery for the measurement of serum C-peptide. The specimens were analyzed at a central laboratory by immunoassay. The 90th percentile for C-peptide for	receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Reference standard used only 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: reference standard used only, no index test used 7) Was the reference standard independent of the index test i.e.

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
of gestational diabetes and obesity with pregnancy outcomes data from the Hyperglycaemia and Adverse Pregnancy Outcome (HAPO) Study Study dates July 2000 to April 2006 Source of funding The study was supported by grants from: The Eunice Kennedy Shriver National Institute of Child Health and Human Development and the National Institute of Diabetes and Digestive and Kidney Diseases	746 (2.9%) were excluded because of glucose unblinding, 1,412 (5.5%) were excluded because they had undergone glucose testing or delivery outside the context of the HAPO Study, and 31 (0.1%) were excluded due to missing key data or improbable results.			the total HAPO cohort (1.7 mg/l) was used to determine the presence of hyperinsulinemia) Entire population No gestational diabetes = 1117/16715 (6.7%) Gestational diabetes = 554/3170 (17.5%) RR (95% CI) = RR 2.62 (2.38 to 2.87) P < 0.00001 Obese women No gestational diabetes = 201/1829 (11%) Gestational diabetes = 168/751 (22.4%) RR (95% CI) = RR 2.04 (1.69 to 2.45) P < 0.00001 Primary ceasarean section (Confirmed from infant birth certificate and defined as the need for the first cesarean delivery at the discretion of the subject's primary obstetrical care provider. Total caesarean deliveries was not used as an outcome because of the various policies regarding delivery at various HAPO Study sites) Entire population No gestational diabetes = 2952/17541 (16.8%) Gestational diabetes =	the index test did not form part of the reference standard: No index test used 8) Was the execution of the index test described in sufficient detail to permit its replication: No index test used 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: No index test used 11) Were the reference standard results interpreted without knowledge of the results of the index test: No index test used 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
The National Centre for Research Resources American Diabetes Association Diabetes UK Kaiser Permenante Medical Centre KK Women's and Children's Centre Mater Mother's Hospital Novo Nordisk The Howard and Carol Bernick Family Foundation				779/3191 (24.4%) RR (95% CI) = RR 1.45 (1.35 to 1.55) P < 0.00001 Obese women No gestational diabetes = 430/1868 (23%) Gestational diabetes = 215/749 (28.7%) RR (95% CI) = RR 1.25 (1.08 to 1.43) P = 0.002 Shoulder dystocia/birth injury (Definition: Additional data were abstracted when either shoulder dystocia or birth injury was suspected. Two members of an outcome review committee (blinded to the mother's glycemic status) reviewed the data to confirm whether either was present.) Entire population No gestational diabetes = 244/19499 (1.3%) Gestational diabetes = 67/3728 (1.8%) RR (95% CI) = RR 1.44 (1.1 to 1.88) P = 0.008 Obese women No gestational diabetes = 32/2252	results reported: No 14) Were withdrawals explained: Yes Other information

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				Gestational diabetes = 26/936 RR (95% CI) = 1.95 (1.17 to 3.26) P = 0.01 * Calculated by NCC-WCH	
Full citation Huynh,J., Ratnaike,S., Bartalotta,C., Permezel,M., Houlihan,C., Challenging the glucose challenge test, Australian and New Zealand Journal of Obstetrics and Gynaecology, 51, 22-25, 2011 Ref Id 154110 Country/ies where the study was carried out Australia Study type Retrospective cohort study Aim of the study	Sample size 8486 women for whom GCT and/or OGTT results were available GCT only = 2291 GCT then OGTT = 416 OGTT only = 5473 Characteristics The baseline characteristics of participants were not presented Inclusion Criteria Women with records for GCT and/or OGTT results on the Austin Pathology database were included Exclusion Criteria Women who were not patients at the Mercy Hospital for Women and those who did not have complete OGTT results were excluded. Where there was more than one OGTT from the same pregnancy, the OGTT furthest away from 26-28 gestational weeks was excluded.	Tests Index test: GCT and FPG. Results for GCT are not presented here because in the published analyses, the majority of women did not receive a 50g glucose load and instead received a 75g glucose load as part of the OGTT. Reference test: 75g OGTT Diagnostic criteria: IADPSG thresholds for gestational diabetes - one or more plasma venous glucose values FPG ≥ 5.1mmol/l, 1 hour ≥ 10.0mmol/l or 2 hour ≥ 8.5mmol/l	Methods 5473 OGTT results were used for the calculation of diagnostic accuracy of FPG and incidence of gestational diabetes interpreted using IADPSG criteria. No details are provided regarding the laboratory methods and standards of glucose testing.	Results Incidence of gestational diabetes Incidence at 24-28 weeks = 1022/5473 (19%) Diagnostic test accuracy of fasting plasma glucose index test at a threshold of ≥ 5.1mmol/l compared with reference standard 2 hour OGTT interpreted using IADPSG criteria thresholds (FPG ≥ 5.1 and/or 1 hour PG ≥ 10.0 mmol/l and/or 2 hour PG ≥ 8.5 mmol/l) at FPG threshold of ≥5.1mmol/l TP: 523* FN: 499* FP: 0* TN: 4451* Sensitivity, % (95% CI): 51.17 (48.11 to 54.23)** Specificity, % (95% CI): 99.99 (99.29 to 100)** LR (95% CI): 4456 (404 to 2391171735)** LR- (95% CI): 0.488 (0.488 to 0.494)** TP - true positive, FN -	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
To estimate how many patients with gestational diabetes would be missed using a glucose challenge test (GCT)/ oral glucose tolerance test (OGTT) combination or a fasting plasma glucose (FPG)/OGTT combination compared to OGTT alone and to assess screening for gestational diabetes using GCT and Australian Diabetes in Pregnancy Society (ADIPS) and International Association of Diabetes in Pregnancy Study Groups (IADPSG) diagnostic				false negative, FP - false positive, TN - true negative * Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html ** 0.5 has been added to each cell (TP, FN, FP, TN) for diagnostic accuracy calculations to take into account the zeros	5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No, the index test did form part of the reference standard 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard

<table>
<tr><th>Bibliographic details</th><th>Participants</th><th>Tests</th><th>Methods</th><th>Outcomes and results</th><th>Comments</th></tr>
<tr><td>criteria

Study dates
May 2005 to April 2007

Source of funding
Not stated</td><td></td><td></td><td></td><td></td><td>results interpreted without knowledge of the results of the index test: Unclear
12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes
13) Were uninterpretable, indeterminate or intermediate test results reported: No
14) Were withdrawals explained: Yes

Other information</td></tr>
<tr><td>Full citation
Kuti,M.A., Abbiyesuku,F.M., Akinlade,K.S., Akinosun,O.M., Adedapo,K.S., Adeleye,J.O., Adesina,O.A., Oral glucose tolerance testing outcomes among women at</td><td>Sample size
765 pregnant women of whom 69 (9%) and 276 (36%) presented in and had data available for the first and second trimesters respectively

Characteristics<table><tr><th></th><th>All</th><th>First trimester</th><th>Second trimester</th><th>Third trimester</th></tr><tr><td>No. of subjects</td><td>765</td><td>69</td><td>276</td><td>420</td></tr><tr><td>Age, years (mean, SD)</td><td>32.3 (4.4)</td><td>31.8 (4.1)</td><td>32.4 (4.5)</td><td>32.4 (4.4)</td></tr><tr><td>Positive family history of</td><td>155 (20.3)</td><td>14 (20.3)</td><td>62 (22.5)</td><td>79 (18.8)</td></tr></table></td><td>Tests
Index test: No index test was used
Reference standard: 2 hour 75 gram oral glucose tolerance test
Diagnostic criteria: WHO 1999 thresholds for gestational diabetes - fasting plasma glucose value (FPG) ≥ 7 mmol/l and/or 2h postload plasma glucose value ≥ 7.8 mmol/l</td><td>Methods
The records of all women referred between June 2007 and July 2009 were reviewed.
For OGTT: Following an overnight fast, two blood samples were taken before and 2h after a 75g of glucose load was administered orally. A diagnosis of gestational diabetes was</td><td>Results
Incidence of gestational diabetes

Incidence of gestational diabetes in the second trimester = 35/276 (12.6%)*

Incidence of gestational diabetes in the second trimester/ Incidence of all gestational diabetes by end of second trimester = 35/47 (74.5%)*

* Calculated by NCC-WCH</td><td>Limitations
NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy
1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes
2) Were selection criteria clearly described: No, exclusion criteria not described</td></tr>
</table>

Bibliographic details	Participants					Tests	Methods	Outcomes and results	Comments
high risk for gestational diabetes mellitus, Journal of Clinical Pathology, 64, 718-721, 2011 Ref Id 153427 Country/ies where the study was carried out Nigeria Study type Retrospective cohort study Aim of the study To determine the prevalence and relationships with known risk factors of gestational diabetes at University College Hospital, Ibadan Study dates June 2007 to July 2009	diabetes, n (%)						made in accordance with the 1999 WHO guidelines. No details regarding standards of laboratory techniques are reported.		3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Reference standard used only 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: reference standard used only, no index test used 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No index test used 8) Was the execution
	History of gestational diabetes, n (%)	14 (1.8)	2 (2.9)	6 (2.2)	6 (1.4)				
	Inclusion Criteria Pregnant women referred to the Metabolic Research Unit (MRU) of University College Hospital, Ibadan for an oral glucose tolerance test. Referrals were made for women at high risk of gestational diabetes based on a history of fetal macrosomia, maternal obesity, previous intrauterine fetal death, first degree relative with diabetes, glycosuria and history of gestational diabetes in a previous pregnancy. Exclusion Criteria Not stated								

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Source of funding Not stated					of the index test described in sufficient detail to permit its replication: No index test used 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: No index test used 11) Were the reference standard results interpreted without knowledge of the results of the index test: No index test used 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: No 14) Were withdrawals explained: There were no withdrawals

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
					Other information
Full citation Senanayake, H., Seneviratne, S., Ariyaratne,H., Wijeratne,S., Screening for gestational diabetes mellitus in southern Asian women, Journal of Obstetrics and Gynaecology Research, 32, 286-291, 2006 Ref Id 181330 Country/ies where the study was carried out Sri Lanka Study type Prospective cohort study Aim of the study To compare fasting plasma glucose	Sample size 271 women referred for oral glucose tolerance testing (OGTT) Characteristics Mean age = 30.7 years (range 17-44) Previous births: First pregnancy n = 90 (34.3%), second pregnancy n = 55 (20.4%), third pregnancy n = 55 (20.3%) Reason for referral: First degree relative with diabetes (52.1%), Maternal age > 35 years (28.1%) Mean gestational age at screening = 26.43 weeks (SD = 5.46) Inclusion Criteria Women with at least one risk factor for gestational diabetes referred to the Reproductive Biology Laboratory of the Faculty of Medicine, University of Colombo for OGTT. Universal screening was not used. Risk factors included having a first degree relative with diabetes, maternal BMI >30kg/cm2 at booking, maternal age > 35 years, previous birth weight > 3.5kg and previous unexplained stillbirth or fetal anomaly. Exclusion Criteria No details are provided	Tests Index test: FPG Reference standard: 2 hour 75 gram oral glucose tolerance test Diagnostic criteria: WHO 1999 thresholds for gestational diabetes - fasting plasma glucose value (FPG) ≥ 7 mmol/l and/or 2h postload plasma glucose value ≥ 7.8 mmol/l	Methods For FPG: The value from the OGTT was used For OGTT: Plasma glucose was estimated using the glucose oxidase method and an automated analyser. No further details are provided.	Results Incidence of gestational diabetes Incidence of gestational diabetes in study population = 75/271 (27.7%) Diagnostic test accuracy of FPG index test at different thresholds compared with reference standard 2 hour OGTT interpreted using WHO 1999 criteria thresholds (FPG ≥ 7.0 or 2 hour PG ≥ 7.8 mmol/l) at FPG test threshold of 4.2 mmol/l TP: 73* FN: 2* FP: 140* TN: 56* Sensitivity, % (95% CI): 97.3 (90.5 to 99.5)* Specificity, % (95% CI): 28.6 (26.0 to 29.4)* LR (95% CI): 1.36 (1.22 to 1.41)* LR- (95% CI): 0.09 (0.02 to 0.36)* at FPG test threshold of 4.4 mmol/l TP: 69* FN: 6* FP: 101* TN: 95* Sensitivity, % (95% CI): 92.0 (83.7 to 96.6)* Specificity, % (95% CI): 48.5 (45.3 to 50.2)* LR (95% CI): 1.78 (1.53 to	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No, exclusion criteria not described 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
(FPG) with postprandial plasma glucose (PPPG) after a carbohydrate meal as screening tests for gestational diabetes in women with one or more risk factors Study dates 1 December 2003 to 31 August 2004 Source of funding None stated				1.94)* LR- (95% CI): 0.16 (0.07 to 0.36)* at FPG test threshold of 4.7 mmol/l TP: 62* FN: 13* FP: 65* TN: 131* Sensitivity, % (95% CI): 82.7 (73.3 to 89.7)* Specificity, % (95% CI): 66.8 (63.2 to 69.5)* LR (95% CI): 2.49 (1.99 to 2.94)* LR- (95% CI): 0.26 (0.15 to 0.42)* at FPG test threshold of 5.0 mmol/l TP: 52* FN: 23* FP: 33* TN: 163* Sensitivity, % (95% CI): 69.3 (59.8 to 77.6)* Specificity, % (95% CI): 83.2 (79.5 to 86.3)* LR (95% CI): 4.12 (2.91 to 5.66)* LR- (95% CI): 0.36 (0.26 to 0.51)* at FPG test threshold of 5.3 mmol/l TP: 34* FN: 41* FP: 16* TN: 180* Sensitivity, % (95% CI): 45.3 (36.7 to 52.7)* Specificity, % (95% CI): 91.8 (88.5 to 94.6)* LR (95% CI): 5.55 (3.20 to	verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No, the index test did form part of the reference standard 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear

<table>
<tr><th>Bibliographic details</th><th>Participants</th><th>Tests</th><th>Methods</th><th>Outcomes and results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td>9.82)*
LR- (95% CI): 0.60 (0.50 to 0.72)*

at FPG test threshold of 7.0 mmol/l
TP: 9* FN: 66* FP: 1* TN: 195*
Sensitivity, % (95% CI): 12.0 (7.3 to 13.3)*
Specificity, % (95% CI): 99.5 (97.7 to 100)*
LR (95% CI): 23.52 (3.18 to 495.46)*
LR- (95% CI): 0.88 (0.87 to 0.95)*

TP - true positive, FN - false negative, FP - false positive, TN - true negative
* Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html</td><td>12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes
13) Were uninterpretable, indeterminate or intermediate test results reported: No
14) Were withdrawals explained: There were no withdrawals

Other information</td></tr>
<tr><td>Full citation
van,Leeuwen M., Opmeer,B.C., Zweers,E.J., van,Ballegooie E., ter Brugge,H.G., de Valk,H.W., Visser,G.H., Mol,B.W., External validation of a clinical</td><td>Sample size
Data from 1301 women included in the previously published cohort study

Characteristics<table><tr><th></th><th></th><th>Gestational diabetes present</th><th>Gestational diabetes not present</th><th>Total</th></tr><tr><td>n</td><td></td><td>47</td><td>1219</td><td>1266</td></tr><tr><td>Age (years)</td><td>≤ 30</td><td>26 (55.3%)</td><td>588 (48.2%)</td><td>614 (48.5%</td></tr></table></td><td>Tests
Index test:
1) Universal screening with 50g 1 hour GCT

2) Application of clinical risk scoring system and 50g 1 hour GCT where indicated
Women at low risk did not receive 50g 1 hour GCT screening</td><td>Methods
Women for whom ethnicity data were not available were excluded from the analysis (35/1301).
All women were screened using a random glucose test (n = 1266) and most women were screened</td><td>Results
Incidence of gestational diabetes
Incidence = 47/1266 = 3.7%

Diagnostic test accuracy of universal 50g 1 hour GCT at 7.8 mmol/l threshold compared with reference standard 2 hour OGTT interpreted using WHO 1999 criteria thresholds (FPG ≥ 7.0 or 2 hour PG ≥</td><td>Limitations
NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy
1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes
2) Were selection</td></tr>
</table>

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
scoring system for the risk of gestational diabetes mellitus, Diabetes Research and Clinical Practice, 85, 96-101, 2009 Ref Id 153872 Country/ies where the study was carried out The Netherlands Study type Prospective cohort study Aim of the study To validate a clinical scoring system to predict gestational diabetes using data from a previously published prospective cohort study	<table><tr><td></td><td>31-34</td><td>7 (14.9%)</td><td>342 (28.1%)</td><td>349 (27.6%</td></tr><tr><td></td><td>≥ 35</td><td>14 (29.8%)</td><td>289 (23.7%)</td><td>303 (23.9%</td></tr><tr><td>BMI (kg/m2)</td><td>≤ 22.0</td><td>8 (17.0%)</td><td>433 (35.5%)</td><td>441 (34.8%</td></tr><tr><td></td><td>22.1 - 25.0</td><td>9 (19.2%)</td><td>398 (32.7%)</td><td>407 (32.2%</td></tr><tr><td></td><td>≥ 25.1</td><td>30 (63.8%)</td><td>388 (31.8%)</td><td>418 (33.0%</td></tr><tr><td>Ethnicity</td><td>Caucasian</td><td>38 (80.9%)</td><td>1094 (89.8%)</td><td>1132 (89.4%</td></tr><tr><td></td><td>Black</td><td>3 (6.3%)</td><td>28 (2.3%)</td><td>31 (2.5%)</td></tr><tr><td></td><td>Asian</td><td>0 (0%)</td><td>5 (0.4%)</td><td>5 (0.4%</td></tr><tr><td></td><td>Other</td><td>6 (12.8%)</td><td>92 (7.5%)</td><td>98 (7.7%)</td></tr></table> Inclusion Criteria Women included in the previously published cohort study that compared the performance of random blood glucose and 50g glucose challenge test as screening tests for gestational diabetes. These women had a singleton pregnancy and received prenatal care from before 24 gestational weeks in two hospitals (in Zwolle and Utrecht) in the Netherlands. Exclusion Criteria Women with a diagnosis of pre-existing type 1 or type 2 diabetes confirmed by a random blood glucose measurement at intake to the study at around gestational week 12.	Women at intermediate risk received 50g 1 hour GCT screening with a threshold of 7.8mmol/l Women at high risk received 50g 1 hour GCT screening with a threshold of 7.1mmol/l Clinical risk scoring system based on age, BMI and race derived by Naylor et al. <table><tr><td>Risk factor</td><td>Sco</td></tr><tr><td>Age (reference category ≤ 30 years)</td><td>0</td></tr><tr><td>31-34 years</td><td>1</td></tr><tr><td>≥ 35 years</td><td>2</td></tr><tr><td>BMI (reference category ≤ 22.0)</td><td>0</td></tr><tr><td>22.1 - 25.0</td><td>2</td></tr><tr><td>≥25.1</td><td>3</td></tr><tr><td>Race (reference category white)</td><td>0</td></tr><tr><td>Black</td><td>0</td></tr><tr><td>Asian</td><td>5</td></tr><tr><td>Other</td><td>2</td></tr></table>	using 50g 1 hour GCT (n = 1246 [98.4%]) at 24-28 gestational weeks. 184 women had at least one abnormal test result and of these 146 (80%) women underwent an OGTT and 38 refused an OGTT. In addition, to estimate the fraction of false negative screening results, women with negative screening results were randomly asked to undergo an OGTT to which 176 consented. Therefore in total 322 women had an OGTT and 46 of these women were diagnosed with gestational diabetes. A multiple imputational procedure was performed to correct for verification bias, to add data for missing OGTT and 50g 1 hour GCT	7.8 mmol/l) TP: 32* FN: 15* FP: 132* TN: 1087* Sensitivity, % (95% CI): 68.1 (53.4to 80.2)** Specificity, % (95% CI): 89.2 (88.6 to 89.6)** LR (95% CI): 6.28 (4.69 to 7.74)** LR- (95% CI): 0.36 (0.22 to 0.57)** Diagnostic test accuracy of selective screening with no 50g 1 hour GCT (low risk) or 50g 1 hour GCT at 7.8 mmol/l threshold (intermediate risk) or 7.1 mmol/l threshold (high risk) compared with reference standard 2 hour OGTT interpreted using WHO 1999 criteria thresholds (FPG ≥ 7.0 or 2 hour PG ≥ 7.8 mmol/l) TP: 30* FN: 17* FP: 153* TN: 1066* Sensitivity, % (95% CI): 63.8 (49.0 to 76.6)** Specificity, % (95% CI): 87.4 (86.9 to 87.9)** LR (95% CI): 5.09 (3.74 to 6.35)** LR- (95% CI): 0.41 (0.27 to 0.59)** TP - true positive, FN - false negative, FP - false positive, TN - true negative	criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: A group selected by screening and a random sample of women not selected by screening were tested using the OGTT reference standard. Data were imputed for other participants 6) Did participants receive the same reference standard regardless of the index test result: A group selected by

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Study dates Not stated Source of funding This study was supported by a grant in the VIDI-program of ZonMW, The Hague and by a grant from Novo Nordisk, Alphen aan den Rijn. The funding sources did not have any involvment in the design, analysis or reporting of the study		Low risk = Clinical risk score 0 or 1 Intermediate risk = Clinical risk score 2 or 3 High risk = Clinical risk score higher than 3 Reference standard: 2 hour 75 gram oral glucose tolerance test Diagnostic criteria: WHO 1999 thresholds for gestational diabetes - fasting plasma glucose value (FPG) ≥ 7 mmol/l and/or 2h postload plasma glucose value ≥ 7.8 mmol/l	results and to add missing BMI and age data. This procedure indicated that 47 women were supposed to be diagnosed with gestational diabetes.	* Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html ** 0.5 has been added to each cell (TP, FN, FP, TN) for diagnostic accuracy calculations to take into account the zeros	index test results received an OGTT. A random sample of women not selected by screening were tested using the OGTT reference standard to correct for verification bias. Data were imputed for other participants 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: Yes 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
					index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: No 14) Were withdrawals explained: Yes Other information

A.14 Diagnostic criteria for gestational diabetes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation Wendland,E.M., Torloni,M.R., Falavigna,M., Trujillo,J., Dode,M.A., Campos,M.A., Duncan,B.B., Schmidt,M.I., Gestational diabetes and pregnancy outcomes - a systematic review of the World Health Organization (WHO) and the International Association of Diabetes in Pregnancy Study Groups (IADPSG) diagnostic	Sample size Nine publications pertaining to eight cohort studies were identified and in total these studies included 44,829 women. Of relevance to this review question are results from two of the included studies, the Brazilian Study of Gestational Diabetes	Tests The relative incidences of several maternal and neonatal outcomes were compared in women with and without gestational diabetes on the basis of diagnosis according to WHO 1999 criteria or IADPSG criteria	Methods Ten electronic databases (MEDLINE, EMBASE, LILACS, the Cochrane Library (CENTRAL), CINAHL, WHO-Afro library, IMSEAR, EMCAT, IMEMR and	Results Eight studies in nine publications were included: Aberg 2001, Black 2010, EBDG 2001, Forsbach 1997, HAPO 2008, HAPO 2010, Khan 1994, Shirazian 2008, Sugaya	Limitations Appendix B: Methodology checklist: systematic reviews and meta-analyses 1) The review addresses an appropriate and clearly focused question that is relevant to the guideline review question: Yes 2) The review collects the type of studies you consider relevant to the guideline review question: Yes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
criteria, BMC Pregnancy and Childbirth, 12 , 2012. Article Number, -, 2012 Ref Id 179445 Country/ies where the study was carried out Brazil Study type Systematic review Aim of the study To summarise the association between gestational diabetes (as defined by World Health Organization (WHO) and International Association of Diabetes and Pregnancy Study Groups (IADPSG) criteria) and adverse pregnancy outcomes in untreated women and evaluate the applicability of the IADPSG criteria beyond the setting of the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study Study dates Searches were run to identify study reports published prior to 15 March 2011 Source of funding Financial support was received from the World Health Organization	(EBDG 2001) and the HAPO study (HAPO 2008) Characteristics Of the eight included studies, one study was performed in the USA, one in Asia, two in the Middle East, one in Europe, two in Latin America (one of which was EBDG 2001) and one was a multi-country study (HAPO 2008). All but one study used venous plasma glucose based on the oral glucose tolerance test (OGTT) to diagnose gestational diabetes EBDG 2001 Ethnicity White 44.9% Mixed 41.4% Black 13.6% Other 0.4% HAPO 2008 Ethnicity White 48.3% Black 11.6% Hispanic 8.5% Asian 29.0% Other 2.6% Inclusion Criteria Prospective or	The WHO 1999 criteria used diagnostic cut points for gestational diabetes that encompassed impaired glucose tolerance and diabetes (fasting plasma glucose ≥ 7 mmol/l ; 2 hour plasma glucose ≥ 7.8 mmol/l) The IADPSG criteria used the following diagnostic cut points for gestational diabetes: a fasting plasma glucose of ≥ 5.1 mmol/l, or a 1 hour result of ≥ 10.0 mmol/l, or a 2 hour result of ≥ 8.5 mmol/l	WPRIM) were searched without language or country restrictions. Classical review articles and reference lists of studies retrieved in full text were also searched for potentially relevant studies. All identified citations were entered into an electronic database and duplicates removed. Two investigators independently screened titles and abstracts of potentially relevant studies. Discrepancies were discussed until consensus was reached Two independent investigators reviewed extracted data using a standardised form. Disagreements were discussed and resolved in a consensus meeting. When	2000 Relative incidence of maternal and neonatal outcomes in women with and without gestational diabetes Caesarean section Data from 2 studies were included EBDG 2001 WHO criteria, women with gestational diabetes = 151/321 IADPSG criteria, women with gestational diabetes = 309/801 Total number of untreated women tested = 4345 HAPO 2008 WHO criteria, women with gestational diabetes = 564/2314 IADPSG criteria, women with gestational	3) The literature search is sufficiently rigorous to identify all the relevant studies: Yes 4) Study quality is assessed and reported: Yes 5) An adequate description of the methodology used is included, and the methods used are appropriate to the question: No : details of data extraction for HAPO 2008 study are inadequate, for the large for gestational age outcome - denominators of the total numbers of women tested for gestational diabetes are different for IADPSG and WHO criteria and the statistical significance of the outcome findings cannot be asessed appropriately for this review question Other information This systematic review investigated a universal screening strategy

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	retrospective cohort studies which included women of any race, parity, age, body weight or other sociodemographic characteristics were considered for inclusion if they provided sufficient information to estimate the associations of the WHO and/or the IADPSG criteria with related perinatal and maternal outcomes Only studies that applied a 2 hour 75 g OGTT performed during the 2nd or the 3rd trimesters universally (in all study participants) and which provided results for a diagnosis based on at least the 2 hour post-load glucose were included. Studies based on capillary glucose measurements were also included Perinatal outcomes examined were large for gestational age births, macrosomia (as defined by the authors) and perinatal mortality (fetal death and early neonatal death). Maternal outcomes that		raw quantitative data were not reported, approximate values were obtained from the figures or calculated from percentages. The methodological quality of included studies was assessed by examining factors that might affect the strength of the association between glucose levels and outcomes. The following factors were assessed in each study: i) adequate selection of participants - consecutive recruitment from antenatal clinics; ii) adequate standardisation of the glucose tolerance test (pre-analytic factors such as anhydrous glucose, plasma immediately separated or kept with glycolytic inhibitors and kept refrigerated until	diabetes = 813/3338 Total number of untreated women tested = 20,732 Large for gestational age (birthweight ≥ 90th centile for gestational age) Data from 2 studies were included EBDG 2001 WHO criteria, women with gestational diabetes = 45/294 Total number of untreated women tested using WHO criteria = 3924 IADPSG criteria, women with gestational diabetes = 87/772 Total number of untreated women tested using IADPSG criteria = 3974 HAPO 2008 WHO criteria, women with gestational diabetes = 361/2642 Total number of	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	were analysed were caesarean delivery and pre-eclampsia (as defined according to individual studies). Only results for women who were untreated were analysed Exclusion Criteria Studies applying the OGTT only in women with certain clinical risk factors (such as family history, obesity, previous gestational diabetes) or in those positive in pre-OGTT glucose screening (with, for example, a 50 g oral glucose challenge test and/or a fasting plasma glucose test) were excluded. Studies that did not distinguish pre-existing diabetes from gestational diabetes, those not allowing the distinction between treated and untreated groups, and those not reporting outcomes for women classified as having a normal OGTT were also excluded In the EBDG 2001 study, the threshold for treatment was 2 hour plasma glucose ≥ 10.0mmol/l, and in the		centrifugation; and analytic factors such as enzymatic method of measurement and laboratory quality control); iii) adequate reporting of losses to follow up; iv) medical staff blinded to OGTT results EBDG 2001 study quality assessment Adequate selection of participants: Yes Adequate test standardisation: Yes Adequate report of losses to follow-up: Yes Medical staff blinded to OGTT results: No HAPO 2008 study quality assessment Adequate selection of participants: Yes Adequate test standardisation: Yes Adequate report of losses to follow-	untreated women tested using WHO criteria = 23,027 IADPSG criteria, women with gestational diabetes = 605/3738 Total number of untreated women tested using IADPSG criteria = 23,217 Perinatal mortality (foetal death and early neonatal death) Data from 1 study were included EBDG 2001 WHO criteria, women with gestational diabetes = 12/330 IADPSG criteria, women with gestational diabetes = 27/802 Total number of untreated women tested = 4431	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	HAPO 2008 study, the thresholds for treatment were fasting plasma glucose > 5.8mmol/l, 2 hour plasma glucose > 10 mmol/l or random plasma glucose ≥ 8.9 mmol/l. Women who were treated were excluded from this systematic review analysis		up: Yes Medical staff blinded to OGTT results: Yes The full database for the EBDG study was available to the authors of the systematic review which permitted analysis for both criteria for all outcomes. Data from the other studies were obtained from published articles cited in the list of references. The EBDG database was used to generate data when results for other studies were not available from the published literature Women who were treated following diagnosis in the EBDG 2001 and HAPO 2008 studies were excluded from the analysis in this systematic review		

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation Jenum,A.K., Morkrid,K., Sletner,L., Vange,S., Torper,J.L., Nakstad,B., Voldner,N., Rognerud-Jensen,O.H., Berntsen,S., Mosdol,A., Skrivarhaug,T., Vardal,M.H., Holme,I., Yajnik,C.S., Birkeland,K.I., Impact of ethnicity on gestational diabetes identified with the WHO and the modified International Association of Diabetes and Pregnancy Study Groups criteria: a population-based cohort study, European Journal of Endocrinology, 166, 317-324, 2012 Ref Id 179806 Country/ies where the study was carried out Norway Study type Prospective cohort study Aim of the study To determine the prevalence of gestational diabetes and its risk factors according to the WHO diagnostic criteria and the modified IADPSG criteria (FPG and 2 hour OGTT values only), to assess the association between ethnic origin and the diagnostic criteria after covariate adjustment, and to discuss the implications of the criteria for public health prevention strategies in a population-based cohort study	Sample size 823 women (74% of those eligible) were included. Of these, data for 759 women were available and included in the analysis Characteristics N = 759 women Mean (standard deviation (SD)) maternal age: 29.9 (4.8) years Parity n (%): Nulliparous 347 (45.7), Uniparous 261 (34.4), Multiparous (≥2) 151 (19.9) Educational level* n (%): <10 years schooling 123 (16.3), Secondary level, 10–12 years schooling 297 (39.5), University/college 333 (44.2) Employed* n (%): 525 (70.0) First-degree relatives with diabetes n (%): 194 (25.6) Mean (SD) gestational week at inclusion: 15 (3.4) Mean (SD) body height: 163.7 (6.7) cm Mean (SD) prepregnancy body mass index (BMI)*: 24.6 (4.8) kg/m2	Tests A 75g OGTT was performed at 28 weeks' gestation after an overnight fast. The reference standard was gestational diabetes diagnosed by applying the WHO 1999 criteria: fasting plasma glucose (FPG) ≥ 7.0 mmol/l or 2 hour plasma glucose (PG) ≥ 7.8 mmol/l The index test was application of the IADPSG criteria, modified as 1 hour plasma glucose values were not available: FPG ≥ 5.1 mmol/l or 2 hour PG ≥ 8.5 mmol/l The WHO 1999 criteria were used for the diagnosis and management of the cases of gestational diabetes during the study. In accordance with the Norwegian national guidelines, women with FPG ≥ 7.0 mmol/l or 2 hour PG ≥ 9.0 mmol/l were referred to secondary care and those with 2 hour PG in the range 7.8–9.0 mmol/l were referred to their general practitioner	Methods The main outcome variable was gestational diabetes. The investigators aimed to enroll at least 800 women, which was expected to result in detection of 100 cases of gestational diabetes Data from questionnaires, anthropometric measurements and venous blood samples drawn after an overnight fast, were collected by specially trained midwives at < 20 and at 28 ± 2 weeks' gestation. The data collected included demographic and socioeconomic factors (education, employment and body height), family history of diabetes, medical and obstetric history and information related	Results Incidence data 99 women (13.0%) were diagnosed with gestational diabetes applying the WHO 1999 criteria (FPG ≥ 7.0 mmol/l and/or 2 hour PG ≥ 7.8 mmol/l) 239 (31.5%) women were diagnosed with gestational diabetes applying the modified IADPSG criteria (FPG ≥ 5.1 mmol/l and/or 2 hour PG ≥ 8.5 mmol/l) Of the 239 women (31.5%) diagnosed with the modified IADPSG criteria: 24.2% were diagnosed exclusively by FPG ≥ 5.1 mmol/l 3.3% were diagnosed exclusively by 2 hour PG ≥ 8.5 mmol/l 4.0% diagnosed by both FPG and	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: Yes 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Study dates Recruitment was between 6 May 2008 to 15 May 2010 Source of funding The Research Council of Norway, the South-Eastern Norway Regional Health Authority, the Norwegian Directorate of Health and collaborative partners in the city of Oslo, Stovner, Grorud and Bjerke administrative districts	*Incomplete data on these variables because of missing values for 6–19 women Inclusion Criteria Women were eligible for inclusion if they satisfied all of the following: a) they lived in the districts b) they planned to give birth at one of the two study hospitals c) they were <20 weeks' gestation d) they could communicate in Norwegian or any other specified languages e) they were able to give written consent to participate Exclusion Criteria Women with pre-existing diabetes or other diseases requiring intensive hospital follow-up during pregnancy were excluded	(GP) after lifestyle advice had been given	to the pregnancy. Body height was measured to the nearest 0.1 cm and body weight was measured to the nearest 0.1 kg. Self-reported prepregnancy bodyweight correlated strongly with weight at inclusion (r=0.97, P<0.001, mean difference 2.0 kg) and was used to calculate prepregnancy BMI	2 hour PG above the cut-off values 492 women were diagnosed with no gestational diabetes (normal glycaemia) applying either WHO 1999 or modified IADPSG criteria: 71 (9.4%) were diagnosed with gestational diabetes applying both the WHO and the modified IADPSG criteria (FPG ≥ 5.1 mmol/l and 2 hour PG ≥ 7.8) 28 (3.7%) were diagnosed with gestational diabetes meeting the WHO 1999 criteria only (FPG < 5.1 mmol/l and 2 hour PG 7.8–8.4 mmol/l) 168 (22.1%) were diagnosed with gestational diabetes meeting the IADPSG criteria only (FPG 5.1–6.9 mmol/l and 2 hour PG < 7.8 mmol/l)	11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: Yes Other information This study investigated a universal screening strategy Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				Diagnostic test accuracy data Diagnostic test accuracy of 2 hour 75g OGTT in the second trimester interpreted using IADPSG thresholds (FPG ≥ 5.1 mmol/l or 2 hour PG ≥8.5 mmol/l for detecting gestational diabetes in the second trimester) compared with reference standard WHO 1999 criteria thresholds (FPG ≥7.0 or 2 hour PG ≥7.8 mmol/l) TP: 71 FN: 28 FP: 168 TN: 492 Sensitivity, % (95% CI): 71.7 (62.4 to 79.7)* Specificity, % (95% CI): 74.5 (73.2 to 75.7)* LR (95% CI): 2.82 (2.32 to 3.28)* LR- (95% CI): 0.38 (0.27 to	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				0.51)* *Diagnostic test accuracy measures and CIs calculated by NCC-WCH technical team based on data reported in the article	
Full citation Kun,A., Tornoczky,J., Tabak,A.G., The prevalence and predictors of gestational diabetes mellitus in Hungary, Hormone and Metabolic Research, 43, 788-793, 2011 Ref Id 181816 Country/ies where the study was carried out Hungary Study type Population-based study Aim of the study To determine the prevalence of gestational diabetes based on the WHO criteria (which were applied at the time of screening) and also the modified IADPSG criteria (the modification was applied because no 1 hour OGTT values were available, only FPG and 2 hour OGTT values Study dates	Sample size n = 1835 of 2260 pregnancies (81.2%) were included in the analysis Characteristics Age < 25 years: 658 25-28 years: 622 29 - 30.9 years: 197 31 - 32.9 years: 139 ≥ 33 years: 219 Pre-pregnancy BMI ≤ 21 kg/m2: 627 21.1.-24.2 kg/m2: 601 24.3 - 26.0: 202 26.1 - 29.1: 197 > 29.1: 208 Previous births 1 : 825 2 : 617 3 : 253 4 : 78 5 : 62	Tests Two definitions of gestational diabetes were used: WHO criteria - gestational diabetes was diagnosed if FPG ≥ 7.0mmol/l or 2 hour plasma glucose value ≥ 7.8mmol/l IADPSG criteria (modified as no 1 hour OGTT samples were drawn) - FPG ≥ 5.1mmol/l or 2 hour plasma glucose ≥ 8.5 mmol/l	Methods A 75g OGTT was performed according to WHO recommendations between 24 and 28 weeks' gestation. Venous blood samples were collected following an overnight fast (≥ 8 hours) and 2 hours after glucose ingestion	Results Incidence data 159/1835 women (8.7%) were diagnosed with gestational diabetes using the WHO criteria 304/1835 (16.6%) were diagnosed with gestational diabetes using the modified IADPSG criteria 104 women were diagnosed with gestational diabetes using both the WHO and IADPSG criteria Diagnostic test accuracy data	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Women who had a pregnancy during the year 2000 were recruited to the study Source of funding Not reported	Inclusion Criteria All pregnant women who lived in Tolna County and gave birth during the year 2000 were included Exclusion Criteria Women were excluded if: a) their pregnancies ended prior to the screening test at 24-28 weeks' gestation b) they had pre-existing diabetes			Diagnostic test accuracy of 2 hour 75g OGTT in the second trimester interpreted using IADPSG thresholds (FPG test ≥ 5.1 mmol/l or 2 hour plasma glucose ≥ 8.5 mmol/l for detecting gestational diabetes in the second trimester) compared with WHO 1999 criteria thresholds (FPG) ≥ 7.0 mmol/l or 2 hour plasma glucose ≥ 7.8 mmol/l)* TP: 104 FN: 55 FP: 200 TN: 1476 Sensitivity, % (95% CI): 65.4 (58.1 to 72.1)* Specificity, % (95% CI): 88.1 (87.4 to 88.7)* LR (95% CI): 5.48 (4.6 to 6.38)* LR- (95% CI): 0.39 (0.31 to 0.48)*	index test did not form part of the reference standard: Yes 8) Was the execution of the index test described in sufficient detail to permit its replication:Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: Yes Other information This study investigated a universal screening strategy Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				*Diagnostic test accuracy measures and CIs calculated by NCC-WCH technical team based on data reported in the article	
Full citation Nallaperumal,S., Bhavadharini,B., Mahalakshmi,M.M., Maheswari,K., Jalaja,R., Moses,A., Anjana,R.M., Deepa,M., Ranjani,H., Mohan,V., Comparison of the world health organization and the International association of diabetes and pregnancy study groups criteria in diagnosing gestational diabetes mellitus in South Indians, Indian Journal of Endocrinology and Metabolism, 17, 906-909, 2013 Ref Id 305955 Country/ies where the study was carried out Chennai, India Study type Retrospective cohort study Aim of the study To compare the IADPSG and WHO criteria to diagnose gestational diabetes in Chennai, India.	Sample size N=1351 pregnant women Characteristics Pregnant women who underwent screening for gestational diabetes at four selected (three private and one government) diabetes centers in Chennai and who, on the basis of elevated glucose levels at screening, were subsequently referred for a 75g OGTT Inclusion Criteria Not stated Exclusion Criteria Not stated	Tests The reference standard was WHO 1999 criteria: fasting plasma glucose (FPG) ≥ 7.0 mmol/l or 2 hour plasma glucose (PG) ≥ 7.8 mmol/l The index test was IADPSG criteria: FPG ≥ 5.1 mmol/l, 1 hour PG ≥ 10.0mmol/l or 2 hour PG ≥ 8.5 mmol/l	Methods All women underwent an oral glucose tolerance test (OGTT) using 75 g glucose load and fasting, 1-h, and 2-h samples were collected.	Results Incidence data 699/1351 women (51.7%) were diagnosed with gestational diabetes applying the WHO 1999 criteria (FPG ≥ 7.0 mmol/l and/or 2 hour PG ≥ 7.8 mmol/l) 699/1351 women (51.7%)were diagnosed with gestational diabetes applying the IADPSG criteria (FPG ≥ 5.1 mmol/l, 1 hour PG ≥ 10.0mmol/l or 2 hour PG ≥ 8.5 mmol/l) Diagnostic test accuracy data Diagnostic test accuracy of 2 hour	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: Yes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Study dates Not stated Source of funding None				75g OGTT in the second trimester interpreted using IADPSG thresholds (FPG ≥ 5.1 mmol/l, 1 hour PG ≥ 10.0mmol/l or 2 hour PG ≥ 8.5 mmol/l for detecting gestational diabetes in the second trimester) compared with reference standard WHO 1999 criteria thresholds (FPG ≥7.0 or 2 hour PG ≥7.8 mmol/l) TP: 559 FN: 140 FP: 140 TN: 512 Sensitivity, % (95% CI): 80 (77.7 to 82.0)* Specificity, % (95% CI): 78.5 (76.1 to 80.8)* LR (95% CI): 3.72 (3.26 to 4.26)* LR- (95% CI): 0.26 (0.22 to 0.29)* *Diagnostic test accuracy measures and CIs calculated by NCC-WCH	8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: None 14) Were withdrawals explained: None Other information This study investigated a selective screening strategy (on the basis of an elevated glucose test at screening) Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				technical team based on data reported in the article	
Full citation Dahanayaka,N.J., Agampodi,S.B., Ranasinghe,O.R., Jayaweera,P.M., Wickramasinghe,W.A., Adhikari,A.N., Chathurani,H.K., Dissanayaka,U.T., Inadequacy of the risk factor based approach to detect gestational diabetes mellitus, Ceylon Medical Journal, 57, 5-9, 2012 Ref Id 182141 Country/ies where the study was carried out Sri Lanka Study type Descriptive Aim of the study To determine the prevalence of gestational diabetes according to the IADPSG criteria and to evaluate a risk factor based approach to diagnosis in Sri Lanka in a cross-sectional study Study dates Not reported Source of funding The Maternal Health Task Force of Gender Health	Sample size n = 405 pregnant women Participant recruitment was performed to cover 10% (n=400) of annual births Characteristics Women participating in the study were from 61 public health midwifery services within three Medical Officer of Health areas Age groups ≤19 n=32 (7.9%) 20-34 n=339 (83.7%) ≥35 n=34 (8.4%) Parity 1 n=171 (42.2%) 2 n=117 (28.9%) 3 n=82 (20.2%) 4 n=27 (6.7%) 5 or more n=8 (2.0%) Gestational age when OGTT was performed 24-28 weeks n=330 (81.5%) 29-32 weeks n=72	Tests The definition of gestational diabetes that was in current use locally was based on risk factors and the WHO definition as follows. History of impaired glucose tolerance (IGT), diabetes, gestational diabetes or polycystic ovary syndrome (PCOS); age > 35 years; weight > 65 kg or BMI > 25 kg/m2; Fundal Height >Predicted Obstetric Average; first-degree relatives with diabetes; birthweight > 3.5 kg in a previous pregnancy; history of unexplained stillbirth or intrauterine death; polyhydramnios or macrosomia; recurrent urinary tract infection; candidiasis; and results of a 75 g OGTT applied according to the WHO criteria (FPG ≥7 mmol/L and/or 2 hour blood	Methods Consenting pregnant women were given verbal and written instructions on preparing for an OGTT and directed to local centres on a day feasible for them. During the visits, venous blood samples were obtained for fasting and at 1 hour and 2 hour post glucose load sugar levels Six trained investigators collected data during the 2-hour waiting period using a pretested interviewer administered questionnaire. Data provided by participants were confirmed using medical records The prevalence of	Results Incidence data Applying the WHO diagnostic criteria FPG only (≥ 7 mmol/l): n=0 (0%) 2 hour glucose only (≥ 7.8 mmol/l): n=28 (6.91%) Both: n=1 (0.25%) Total: n=29 (7.16%) Applying the IADPSG diagnostic criteria FPG only (≥ 5.1 mmol/l): n=19 (4.69%) 1 hour glucose only (≥ 10.0 mmol/l): n=0 (0%) 2 hour glucose only (≥ 8.5 mmol/l): n=3 (0.74%) FPG and 1 hour value: n=4 (0.99%) FPG and 2 hour	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes, inclusion criteria described, exclusion criteria were not reported however the study is of cross sectional design 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: Yes 8) Was the execution of the index test described in sufficient detail to

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	(17.8%) >32 weeks n=3 (0.7%) Gestational age at registration ≤8 weeks n=232 (57.3%) 9-12 weeks n=134 (33.1%) >12 weeks n=39 (9.6%) Inclusion Criteria All pregnant women from the three participating areas at more than 24 weeks' gestation but not more than 28 weeks' gestation were invited to participate Exclusion Criteria Not reported	glucose ≥ 7.8 mmol/l) Gestational diabetes defined using the IADPSG criteria was as follows. Any woman with one or more of the following results in a 75g OGTT: FPG ≥ 5.1 mmol/l , 1 hour plasma glucose ≥ 10 mmol/l or 2 hour plasma glucose ≥ 8.5 mmol/l	gestational diabetes was determined using the WHO and IADPSG criteria separately. Prevalence was defined as the percentage of women who had gestational diabetes according to at least one set of criteria. Women with risk factors were selected to establish the percentage of women that could have been diagnosed if current local recommendations were followed. Risk factors from previous pregnancies and the current pregnancy were included as well as risk factors and early indicators of gestational diabetes. Women with a single risk factor for gestational diabetes were then examined and classified using the	value: n=0 (0%) 1 hour and 2 hour values: n=7 (1.73%) FPG, 1 hour and 2 hour values: n=3 (0.74%) Total n=36 (8.89%) Diagnostic test accuracy data Diagnostic test accuracy of 2 hour 75g OGTT in the second trimester interpreted using the IADPSG criteria (FPG ≥ 5.1 mmol/l or 1 hour plasma glucose ≥ 10 mmol/l or 2 hour plasma glucose ≥ 8.5 mmol/l for detecting gestational diabetes in the second trimester) compared with the WHO 1999 criteria (FPG ≥ 7.0 mmol/l or 2 hour plasma glucose ≥ 7.8 mmol/l)* TP: 22 FN: 14 FP: 7 TN: 0	permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: No 14) Were withdrawals explained: Yes Other information The population in this study was from Sri Lanka. Being of South Asian descent is an independent risk factor for developing gestational diabetes as South Asian populations have a high prevalence of diabetes -Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
			WHO criteria. The proportion of women diagnosed with gestational diabetes was then compared with the recommended IADPSG guidelines	Sensitivity, % (95% CI): 60.8 (59.5 to 68.8)** Specificity, % (95% CI): 6.2 (0.32 to 36.9)** LR (95% CI): 0.65 (0.6 to 1.21)** LR- (95% CI): 6.27 (0.72 to 3400786)** Diagnostic test accuracy of screening with FPG (IADPSG) in the first trimester using the IADPSG criteria (FPG ≥ 5.1 mmol/l) versus second trimester 2 hour 75g OGTT using the WHO 1999 criteria(FPG ≥ 7.0 mmol/l or 2 hour plasma glucose ≥ 7.8 mmol/l)* Retrospective analysis of 16/400 women screened with FPG during the first trimester TP : 0 FN : 3 FP : 2 TN : 11 Sensitivity % (95% CI) : 12.5	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				(0.63 to 60.2)** Specificity % (95% CI) : 82.1 (78.6 to 94.7)** LR (95% CI) : 0.7 (0.0 to 10.61)** LR- (95% CI) : 1.07 (0.46 to 1.27)** *Diagnostic test accuracy measures and CIs calculated by NCC-WCH technical team based on data reported in the article **0.5 has been added to each cell (TP, FN, FP, TN) for diagnostic accuracy calculations to take zeros into account	

A.15 Interventions for gestational diabetes: GDM – interventions

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation	Sample size	Interventions	Details	Results	Limitations

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Asemi,Z., Samimi,M., Tabassi,Z., Esmaillzadeh,A., The effect of DASH diet on pregnancy outcomes in gestational diabetes: A randomized controlled clinical trial, European Journal of Clinical Nutrition, 68, 490-495, 2014 Ref Id 318940 Country/ies where the study was carried out Iran Study type Randomised controlled trials Aim of the study To investigate the effects of the DASH (Dietary Approahes to Stop Hypertension) eating plan on outcomes in pregnant women with gestational diabetes	n=52 (26 in each arm) Characteristics Maternal age (years) DASH group = 31.9 ± 6.1 Control group = 30.7 ± 6.3 p = 0.47 Prepregnancy weight (kg) DASH group = 68.8 ± 10.9 Control group = 160.4 ± 6.4 p = 0.11 Weight at baseline (kg) DASH group = 74.7 ± 10.7 Control group = 79.7 ± 11.8 p = 0.11 Prepregnancy BMI (kg/m2) DASH group = 26.9 ± 3.4 Control group = 28.8 ± 4.8 p = 0.11 Gestational age before intervention (wks) DASH group = 25.8 ± 1.4 Control group = 25.9 ± 1.4 p = 0.77 Inclusion criteria Primigravida pregnant women aged 18-40 years, screened with 50g Glucose Challenge Test and those with results >140mg/dl underwent diagnostic testing for gestational diabetes by 100g OGTT at 24-28 gestational weeks. Exclusion criteria Women with a previous glucose intolrance/gestational diabetes diagnosis, premature preterm rupture of membranes, placenta abruption, preeclampsia, need for insulin during the intervention, complete bed rest, hypothyroidism, urinary tract infection,	DASH diet: similar to the control diet, but was rich in fruits, vegetables, whole grains and low-fat dairy products and low in saturated fats, cholesterol, refined grains and sweets Control diet: 45-55% carbohydrates, 15-20% protein and 25-30% total fat	After stratification for BMI and weeks of gestation (< 26 weeks or ≥26 weeks), women were randomly assigned (using computer-generated random numbers) to treatment groups for 4 week intervemtion. Women were asked not to change their physical activity, as well as not to take any antihyperglycaemic or lipid-lowering medications. Compliance with diets was assessed once a week with telephone calls. Participants completed three 1 day dietary records (2 weekdays and 1 weekend day) throughout the study which were assessed using Nutritionist IV softwarw modified for Iranian foods to obtain nutrient intake. Statistical analysis Power calculation (based on mean birth weight) estimated that 21 participants per groups were necessary	Caesarean section DASH diet group = 12/26 (46.2%) Control diet group = 21/26 (80.8%) p = 0.01	NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Yes A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Yes A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes B. Performance bias B1: The comparison groups

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Study dates January 2012 to June 2013 Source of funding Vice-Chancellor for Research, KUMS, Kashan, Iran	smoking and kidney or liver diseases as well as those taking oestrogen therapy.				received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. No B3: Individuals administering care were kept 'blind' to treatment allocation. Unclear C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes C2: a. How many participants did not complete treatment in each group? None b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					treatment). Yes C3: a. For how many participants in each group were no outcome data available? None b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Yes D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear

Other information
None</td></tr>
<tr><td>Full citation
Avery,M.D., Leon,A.S., Kopher,R.A., Effects of a partially home-based exercise program for women with gestational diabetes, Obstetrics and Gynecology, 89, 10-15, 1997
Ref Id
177086
Country/ies where the study was carried out
United States of America
Study type
Randomised controlled trial.

Aim of the study
To test the</td><td>Sample size
Total sample size, after exclusions and attrition, comprised 29 women (15 intervention, 14 control).

Characteristics
<table><tr><th>Characteristic</th><th>Intervention</th><th>Control</th><th>P-value</th></tr><tr><td>Mean gestation at diagnosis</td><td>28.7 ± 3.0</td><td>26.3 ± 8.1</td><td>0.30</td></tr><tr><td>Mean 3 hour OGTT, mg/dl</td><td></td><td></td><td></td></tr><tr><td>Fasting</td><td>85 ± 6.8</td><td>84 ± 0.28</td><td>0.84</td></tr><tr><td>1 hour</td><td>191 ± 24.7</td><td>203 ± 39.6</td><td>0.39</td></tr><tr><td>2 hours</td><td>185 ± 18.8</td><td>187 ± 25.8</td><td>0.86</td></tr><tr><td>3 hours</td><td>151 ± 28.2</td><td>138 ± 49.2</td><td>0.44</td></tr><tr><td>Parity, mean</td><td>1.5</td><td>0.4</td><td>0.005</td></tr></table></td><td>Interventions
Intervention
30 minutes exercise three to four times per week until delivery.

Control
Maintained usual physical activity level.</td><td>Details
Subjects were not blinded to the intervention.

All women diagnosed with GDM who met eligibility crtieria were invited to participate.

Following diagnosis with GDM eligible women were randomised to either treatment group using block randomisation of numbers from random number tables.

Intervention group participants undertook 30 minutes of exercise three or four times per week. The exercise comprised 5 minutes warm-up and cool-down before and after a 20 minute work out. Exercise intensity was 70% of the age-</td><td>Results
Caesarean delivery
Treatment: 3/15
Control: 3/14
RR = 0.93 (95% CI 0.22 to 3.87)*

Macrosomia (> 4000g)
Treatment: 3/15
Control: 3/14
RR = 0.93 (95% CI 0.22 to 3.87)*

Neonatal hypoglycaemia
Treatment: 0/15
Control: 0/14
RR not calculable.

Requirement for insulin
Treatment: 4/15
Control: 2/14
RR = 1.86 (95% CI 0.40 to 8.62)*

*Calculated by the NCC-WCH technical team.</td><td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Yes

A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
effectiveness of a program of moderate-intensity exercise on blood glucose control of women with gestational diabetes mellitus. Study dates Not reported. Source of funding Not reported.	<table><tr><td>Caucasian, n/N</td><td>15/15</td><td>12/14</td><td>0.22</td></tr></table> Inclusion criteria Physician or nurse-midwife certified diagnosis of GDM, ≤ 34 weeks' gestation, ability to read and write English, no other important medical or obstetric complications, aged 18 to 40 years, no current regular exercise regimen similar to the intervention. Diagnosis criteria for 3 hour OGTT were based on the O'Sullivan and Mahan criteria (1964): Fasting < 5.0mmol/l 1 hour < 9.2mmol/l 2 hours < 8.1mmol/l 3 hours < 6.9mmol/l Exclusion criteria Not reported.		related maximum (0.70 x (220 - age in years)). Two exercise sessions per week were monitored by study staff. Unsupervised exercise primarily involved walking. Three women used a cycle ergometer. Control subjects continued diet therapy and their usual physical activity level. Women were asked not to change their current physical activity level. All subjects recorded fasting and 2 hour post-prandial glucose levels three days per week. Insulin therapy was initiated if required and recorded during data collection. Dietary intake was assessed using a food frequency questionnaire. Neonatal hypoglycaemia was defined as blood glucose < 45mg/dl 3 or 5 hours after birth.		or treatment allocation). Unclear A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes overall though parity differed between groups. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. No B3: Individuals administering care were kept 'blind' to treatment allocation. Unclear C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			Statistical analysis Most data were analysed using Student's t-tests, paired or unpaired for within- and between-group differences. X2 tests, Fisher's exact tests and Mann-Whitney U tests were used to analyse nominal or ordinal data, where appropriate. Results were considered significant for p-values < 0.05.		adjusted to allow for differences in length of follow-up). Yes - analyses incorporated a time element (regression). C2: a. How many participants did not complete treatment in each group? 1 in the intervention group, 3 controls. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Unclear C3: a. For how many participants in each group were no outcome data available? None for the relevant outcomes for this review.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Yes D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>were kept 'blind' to other important confounding and prognostic factors. Unclear

Other information
None.</td></tr>
<tr><td>Full citation
Bertini,A.M., Silva,J.C., Taborda,W., Becker,F., Lemos Bebber,F.R., Zucco Viesi,J.M., Aquim,G., Engel,Ribeiro T., Perinatal outcomes and the use of oral hypoglycemic agents, Journal of Perinatal Medicine, 33, 519-523, 2005
Ref Id
177112
Country/ies where the study was carried out
Brazil
Study type
Open label randomised</td><td>Sample size
70 women randomised to treatment with insulin (n=27), glibenclamide (n=24) and acarbose (n=19)

Characteristics
<table><tr><th>Mean ± SD</th><th>Glibenclamide (n=24)</th><th>Insulin (n=27)</th><th>p value</th></tr><tr><td>Age at start of treatment (years)</td><td>31.2 ± 4.5</td><td>28.7 ± 6.0</td><td>NS</td></tr><tr><td>Number of pregnancies</td><td>3.2 ± 6.5</td><td>2.5 ± 1.6</td><td>NS</td></tr><tr><td>BMI</td><td>27.5 ± 5.8</td><td>27.0 ± 7.2</td><td>NS</td></tr><tr><td>Weight gain</td><td>10 ± 5.2</td><td>11.5 ± 3.8</td><td>NS</td></tr></table>
Inclusion criteria
Women attending a multidisciplinary maternity unit in Joinville who were diagnosed using 2 h</td><td>Interventions
Glibenclamide : An initial dose of 5mg in the morning was increased every week as necessary to a maximum dose of 20mg/day. Blood glucose was reviewed in clinic weekly.
Insulin : Women were admitted to hospital for 24 hrs to learn how to use insulin and to receive guidance. Insulin was started at a dosage of 0.7 units of insulin/kg actual body weight, increasing by 0.1 IU/kg in each trimester. Rapid action and</td><td>Details
Women had three days of diet and physical activity and then their fasting and postprandial glucose levels were measured. Acceptable levels for FPG were 90mg/dl and postprandial tests 100mg/dl.Participation in the trial was offered to those who did not meet these thresholds. No details of diet or exercise are given.

Blood glucose was reviewed in clinic weekly. Women were tested in the fasting state and 2 hours after breakfast. If either test was abnormal, testing was performed after lunch and dinner to estblish glucose profile and adjust doses as necessary.</td><td>Results
Caesarean Section
Glibenclamide = 12/24 (50%)
Insulin = 12*/27 (44.4%)
Treatment Failure
Glibenclamide = 5/24 (20.8%)

Large for gestational age (defined as >90th percentile by growth curves)
Glibenclamide = 6/24 (25%)
Insulin = 1/27 (3.7%)
Neonatal hypoglycaemia (defined as <40mg/dl, both treatments interrupted 14-24 hours prior to delivery)
Glibenclamide = 8/24 (33.3%) (1 NICU admission, 7 managed with maternal milk)
Insulin = 1/27 (3.7%) (1 managed with maternal milk)
NICU admission
Glibenclamide = 1/24 (delivered at 36 GW, admitted for 2 days)
Insulin = 0/27
Birth injuries (no definition)
Glibenclamide = 0/24</td><td>Limitations
NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials
Appropriate randomisation method: Yes
Adequate allocation concealment: Yes
Groups comparable at baseline: Yes
Groups received the same care (apart from the intervention): Yes
Participants kept 'blind' to allocation: No
Care givers kept 'blind' to allocation: No
Follow up equal for groups: Yes
How many participants did not</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
controlled trial Aim of the study To compare neonatal outcomes from women with gestational diabetes who were randomised to treatment with insulin, glibenclamide or acarbose Study dates 1 October 2003 to 1 July 2004 Source of funding Not stated	OGTT and WHO diagnostic criteria/local Health Ministry interpretation of this (FPG ≥ 110mg/dl and 2hr value ≥ 140mg/dl). Women also had to have given informed consent and been from 11 to 33 gestational weeks of a singleton pregnancy at diagnosis. Exclusion criteria Women with concomitant pathologies that would affect treatment or perinatal results were excluded as were any women who wished to discontinue or who the researchers believed required faster glucose control (eg corticoid therapy).	slow acting insulins were used in equal doses before main meals and at bedtime respectively.	Treatment failure was defined taking the maximum dose without achieving glucose control. Oral medication was stopped in treatment failure and insulin therapy started. Statistical analysis ANOVA was performed using Excel with a 95% significance threshold.	Insulin = 0/27 Neonatal death Glibenclamide = 0/24 Insulin = 0/27	complete treatment in each group?: 1 woman from an unknown group Were the groups were comparable for treatment completion: For how many participants in each group were no outcome data available?: Yes The groups were comparable with respect to the availability of outcome data: Yes Appropriate length of follow-up: Yes Precise outcome definitions used: Yes Outcome determined using valid and reliable methods: Yes Investigators kept 'blind' to allocation: Unclear Investigators kept 'blind' to other important confounding and prognostic factors: Unclear Other information None.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation Bevier,W.C., Fischer,R., Jovanovic,L., Treatment of women with an abnormal glucose challenge test (but a normal oral glucose tolerance test) decreases the prevalence of macrosomia, American Journal of Perinatology, 16, 269-275, 1999 Ref Id 177114 Country/ies where the study was carried out United States of America Study type Randomised controlled trial. Aim of the study To examine the effectiveness of a treatment regimen in reducing foetal macrosomia,	Sample size The total sample size comprised 103 women, 83 of whom were included in final analyses (35 intervention, 48 control). Characteristics Characteristic / Treatment / Control Mean age, years: 27.4 ± 5.4 / 26.3 ± 6.0 Mean number gravida: 2.8 ± 1.7 / 3.1 ± 2.0 Mean number parous: 1.3 ± 1.5 / 1.6 ± 1.7 Mean weight at 28 to 30 weeks, lbs: 150.4 ± 25.2 / 159.7 ± 26.5 Ethnicity, n (%) White, non-Hispanic: 2 (4%) / 2 (6%) White, Hispanic: 45 (94%) / 33 (94%) African-American: 1 (2%) / 0 (0%) P-values not reported. Inclusion criteria Positive oral challenge test screening result with a negative oral glucose tolerance test. Thresholds for diagnosis were not reported.	Interventions Intervention Dietary counselling Instruction in self monitoring of blood glucose 30kcal/kg/day or 24kcal/kg/day if body weight > 120% of ideal Control No diet or self monitoring of blood glucose	Details Women with a positive oral challenge test but a negative oral glucose tolerance test were randomly assigned to each treatment arm. Self monitoring blood glucose diaries were reviewed weekly by a clinic nurse. Random blood glucose measures were also monitored in the clinic. Insulin therapy was initiated if fasting blood glucose > 90mg/dl (5.0mmol/l) or one hour post-prandial glucose > 120mg/dl (6.7mmol/l) on three or more occasions. Random blood glucose checks were performed on controls and insulin started if glucose > 120mg/dl (6.7mmol/l). Birth weight was recorded in grams and as a percentile using gender and ethnicity-specific curves. Shoulder dystocia was not defined. Statistical analysis	Results Mode of delivery Vaginal spontaneous Treatment: 22/35 Control: 30/48 RR = 1.02 (95% CI 0.73 to 1.43)* Vaginal induced Treatment: 6/35 Control: 0/48 RR = 21.37 (95% CI 1.24 to 367.31)* Vaginal forceps Treatment: 0/35 Control: 1/48 RR = 0.45 (95% CI 0.02 to 10.73)* Vaginal vacuum Treatment: 2/35 Control: 1/48 RR = 2.84 (95% CI 0.27 to 30.10)* Primary caesarean Treatment: 3/35 Control: 3/48 RR = 1.41 (95% CI 0.30 to 6.58)* Repeat caesarean Treatment: 2/35	Limitations NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Unclear A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear A3: The groups were comparable at baseline, including all major confounding and prognostic factors.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
maternal and infant morbidity, maternal complications and operative delivery in women with an abnormal glucose challenge test but a normal oral glucose tolerance test. Study dates Not reported. Source of funding Not reported.	Exclusion criteria Evidence of hypertension, collagen disease, chronic renal disease, cardiac or pulmonary disease, rhesus sensitisation, a history of pre-term labour or small for gestational age deliveries.		Analyses included X2 tests for categorical variables or Student's t-tests for continuous variables.	Control: 9/48 RR = 0.26 (95% CI 0.06 to 1.13)* Large for gestational age Treatment: 1/35 Control: 12/48 RR = 0.09 (95% CI 0.01 to 0.66)* Shoulder dystocia Treatment: 1/35 Control: 2/48 RR = 0.68 (95% CI 0.06 to 7.21)*	Unclear B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Unclear B2: Participants receiving care were kept 'blind' to treatment allocation. Unclear B3: Individuals administering care were kept 'blind' to treatment allocation. Unclear C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? Unclear - 83 out of

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					103 participants included in final analyses (48 control and 35 intervention). b. The groups were comparable for treatment completion. Unclear C3: a. For how many participants in each group were no outcome data available? Unclear - 83 out of 103 participants included in final analyses (48 control and 35 intervention). b. The groups were comparable with respect to the availability of outcome data. Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Unclear D2: The study used

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					a precise definition of outcome. No - shoulder dystocia not defined. D3: A valid and reliable method was used to determine the outcome. Unclear D4: Investigators were kept 'blind' to participants' exposure to the intervention. No D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information None.
Full citation Bonomo,M., Corica,D., Mion,E., Goncalves,D., Mottat,G., Merati,R., Ragusa,A., Morabito,A.,	Sample size Total sample size comprised 300 women (150 intervention, 150 no treatment, 150 control). Characteristics Characteristic / No treatment / Diet / Control	Interventions Intervention Dietary advice of 24 to 30kcal/hr/day based on pre-pregnancy weight (50 to 50% carbohydrates, 25	Details After diagnosis eligible women were stratified by age and BMI then randomly assigned to either diet or no treatment using random number tables. Control subjects were	Results Large for gestational age Diet: 9/150 No treatment: 21/150 RR = 0.43 (95% CI 0.20 to 0.91)* Hypoglycaemia	Limitations NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual A. Selection bias

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Evaluating the therapeutic approach in pregnancies complicated by borderline glucose intolerance: A randomized clinical trial, Diabetic Medicine, 22, 1536-1541, 2005 Ref Id 177122 Country/ies where the study was carried out Italy Study type Randomised controlled trial. Aim of the study To determine whether an appropriate diet could reduce the prevalence of macrosomia in women with mild gestational diabetes. Study dates 1997 to 2002.	<table><tr><td>Mean age, years</td><td>30.7 ± 5.1</td><td>31.1 ± 4.7</td><td>31.1 ± 4.4</td></tr><tr><td>Primiparous, %</td><td>42.0</td><td>45.3</td><td>40.0</td></tr><tr><td>Body mass index, kg/m2</td><td>23.0 ± 4.5</td><td>23.1 ± 4.4</td><td>23.0 ± 4.1</td></tr><tr><td>Fasting plasma glucose, mmol/l</td><td>4.77 ± 0.52</td><td>4.68 ± 0.45</td><td>4.56 ± 0.40</td></tr></table> No significant differences between groups were observed. No p-values were reported. Inclusion criteria Women of Caucasian origin, women with a positive 50g GCT but negative 100g OGTT. Criteria for the glucose tests were: 1 hour GCT 7.8mmol/l OGTT: 0 hours 5.3mmol/l, 1 hour 10.1mmol/l, 2 hours 8.7mmol/l and 3 hours 7.8mmol/l Exclusion criteria Women with a normal GCT (except control subjects), one abnormal OGTT value and women fulfilling criteria for full GDM.	to 30% protein, 20 to 25% fat). No treatment No special care, diet or pharmacological intervention. Control Normal GCT women matched by strata of age and BMI.	matched according to these strata. Women assigned to the diet group were evaluated every 2 weeks as out-patients. Dietary habits and compliance were discussed and fasting 2 hour post-prandial glucose measurements taken. Glucose targets were 5.0mmol/l fasting and < 6.7mmol/l at 2 hours. 25 women refused to participate before randomisation. Recruitment was continued until n = 300 women were enrolled. After randomisation 21 women were replaced as they left care (6 women in the diet group) or were diagnosed with GDM (9 in the diet group, 6 in the no treatment group). The study was not blinded. Outcomes included: LGA (> 90th percentile for gestational age) Hypoglycaemia (any two blood glucose values < 1.7mmol/l)	Diet: 5/150 No treatment: 6/150 RR = 0.83 (95% CI 0.26 to 2.66)* Neonatal unit stay Diet: 5/150 No treatment: 7/150 RR = 0.71 (95% CI 0.23 to 2.19)* Caesarean section Diet: 44/150 No treatment: 42/150 RR = 1.05 (95% CI 0.73 to 1.50)* *Calculated by the NCC-WCH technical team.	A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Yes A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). No A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. No

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Source of funding Not reported.			Rate of caesarean sections Rate of admission to NICU Statistical analysis Sample size was calculated to provide 80% power at a significance level of 0.05 to detect an 11% change in LGA rates between groups. Differences in means were assessed using Student's t-tests, ANOVA or Scheffe's tests. Categorical data were assessed using Yates' corrected X2 tests. Kruskal-Wallis tests were used to compare medians.		B2: Participants receiving care were kept 'blind' to treatment allocation. No B3: Individuals administering care were kept 'blind' to treatment allocation. No C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? None - though women were replaced (15 in the diet group, 6 in the no treatment group). b. The groups were comparable for treatment completion (that is, there were no

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					important or systematic differences between groups in terms of those who did not complete treatment). No - 6 women in the diet group left care, none left for this reason in the no treatment group. C3: a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. No D5: Investigators were kept 'blind' to other important confounding and prognostic factors. No Other information None.
Full citation Brankston,G.N., Mitchell,B.F., Ryan,E.A., Okun,N.B., Resistance exercise decreases the	Sample size Total sample size comprised 32 women (16 intervention, 16 control). Characteristics	Interventions Control Standard diabetic diet (40% carbohydrate, 40% protein, 20% fat) comprising 24 to 30kcal/kg/day	Details Following diagnosis with GDM eligible women were randomised to either treatment group using random number tables. Allocation was concealed using opaque	Results Requirement for insulin therapy Intervention: 7/16 Control: 9/16 RR = 0.78 (95% CI 0.39 to 1.58)*	Limitations NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
need for insulin in overweight women with gestational diabetes mellitus, American Journal of Obstetrics and Gynecology, 190, 188-193, 2004 Ref Id 177127 Country/ies where the study was carried out Canada Study type Randomised controlled trial. Aim of the study To determine the effect of circuit-type resistance training on the requirement for insulin in women with gestational diabetes mellitus. Study dates Not reported. Source of	<table><tr><th>Characteristic</th><th>Control</th><th>Intervention</th><th>P-value</th></tr><tr><td>Mean maternal age, years</td><td>31.3 ± 5.0</td><td>30.5 ± 4.4</td><td>0.63</td></tr><tr><td>Mean pre-pregnant BMI, kg/m2</td><td>28.0 ± 5.7</td><td>25.9 ± 3.4</td><td>0.21</td></tr><tr><td>Mean, gestation at first clinic visit</td><td>29.6 ± 2.1</td><td>29.0 ± 2.0</td><td>0.44</td></tr></table> Inclusion criteria Aged 20 to 40 years, gestational age between 26 and 32 weeks, BMI below 40kg/m2, non-smokers, not involved in a regular exercise program. GDM was diagnosed using a screening test followed by an OGTT. Screening test diagnostic criteria: 1 hour glucose ≥ 10.3mmol/l (185mg/dl) OGTT diagnostic criteria required that two or more of the following values be exceeded: Fasting ≥ 5.3mmol/l (95mg/dl) 1 hour ≥ 10.6mmol/l (191mg/dl) 2 hours ≥ 8.9mmol/l (160mg/dl) Exclusion criteria Not reported.	of ideal pre-pregnancy body weight. Intervention Diet as per the control group plus a progressive physical activity program of circuit-type exercise.	sequentially numbered envelopes. Women in the control group were asked not to begin a structured exercise program before delivery. Intervention group participants were instructed to exercise three times per week. Exercise was circuit-based with up to a minute's rest between each exercise. Resistance was provided using resistance bands. Women were instructed to exercise so that it felt "somewhat hard". As exercises became easier, difficulty was increases so as to maintain intensity. The number of sets and repetitions increased over the course of four weeks. Subjects monitored their own heart rate to ensure it was not above 140 beats/minute. Insulin therapy was initiated if the following values were consistently exceeded during treatment: Fasting ≥ 5.3mmol/l	Data for birth weight and caesarean delivery were not reported. No significant differences were observed between groups. *Calculated by the NCC-WCH technical team.	A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Yes A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Yes A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes but ethnicity not reported. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s)

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
funding Not reported.			(95mg/dl) 1 hour ≥ 7.8mmol/l (140mg/dl) 2 hours ≥ 6.7mmol/l (120mg/dl) The main outcome was the requirement for insulin in women. Neonatal outcomes also included birth weight. Statistical analysis Sample size was calculated to provide 80% power to detect a 25% difference in insulin use at the 0.05 significance level. Ideal sample size was 32 participants in total. X2 tests were used to analyse between-group differences for categorical variables. Independent sample t-tests were used to analyse continuous variables. Variables that were not normally distributed were analysed using Mann-Whitney U tests.		studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? Unclear. One woman dropped out, group not reported. Two in the intervention group did not start the exercise program. Three were advised against exercise by

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					physicians, group not reported. 32/38 enrolled completed the study. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Unclear C3: a. For how many participants in each group were no outcome data available? None for outcomes relevant to this review. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					available). Yes D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. Yes, thresholds for insulin therapy were reported. D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information None.

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr>
<td>Full citation
Coustan,D.R., Lewis,S.B., Insulin therapy for gestational diabetes, Obstetrics and Gynecology, 51, 306-310, 1978
Ref Id
177185
Country/ies where the study was carried out
United States of America
Study type
Partially randomised trial.

Aim of the study
To compare the effect of treatment with diet plus insulin versus diet alone and versus neither diet nor insulin on birthweight in women with gestational diabetes mellitus.</td>
<td>Sample size
The total sample size comprised 72 women (27 diet + insulin, 11 diet alone, 34 control).

Characteristics
<table>
<tr><td></td><td>Control</td><td>Insulin+diet</td><td>Diet alone</td></tr>
<tr><td>Pregnancy weight, lb</td><td>148.5 ± 45.6</td><td>150.1 ± 40.1</td><td>158.3 ± 55.6</td></tr>
<tr><td>Weeks in study</td><td>5.4</td><td>7.9</td><td>6.1</td></tr>
</table>
No p-values were reported.

Inclusion criteria
Women were given an OGTT if they had: family history of diabetes, a previous baby weighing more than 8.5lb (3.864kg), poor obstetric history or glucosuria at any prenatal visit.

Cut-offs for the OGTT, modified for serum glucose, were < 95mg/dl for fasting values, < 180mg/dl at 1 hour, < 160mg/dl at 2 hours and < 135mg/dl at 3 hours. GDM was diagnosed if two or more glucose test results met or exceeded these values.

Exclusion criteria
Not reported.</td>
<td>Interventions
Diet alone
A diet of 30-35kcal/kg ideal weight/day comprising 500kcal protein with the rest of the intake split equally between fat and carbohydrates.

Diet + insulin
Diet plus 20 units NPH insulin and 10 units regular insulin 30 minutes before breakfast.

Control
Dietary counselling as per standard prenatal care with 90g protein and 15 to 25lb weight gain recommended.</td>
<td>Details
Following diagnosis of GDM women were enrolled into the study. The first 20 women were not randomised: 10 were diagnosed < 36 weeks' gestation and were assigned to the intervention group, 10 were diagnosed > 36 weeks' gestation and were assigned to the control group. Treatment was started immediately following diagnosis.

Subjects were evaluated every two weeks by taking fasting glucose measurements and 2 hour post-prandial measurements after breakfast. After 34 week's gestation women were seen weekly.

Diet and insulin therapy were stopped on the day of delivery.

Outcomes included:
Perinatal mortality
Shoulder dystocia
Macrosomia
Caesarean delivery</td>
<td>Results
Macrosomia (neonates > 3.864kg)
Diet + insulin vs. diet alone
Diet + insulin: 2/27
Diet alone: 4/11
RR = 0.20 (95% CI 0.007 to 5.66)*

Diet alone vs. no diet
Diet alone: 4/11
No diet: 17/34
RR = 0.72 (95% CI 0.31 to 1.69)*

Total number of caesarean sections
Diet + insulin vs. diet alone
Diet + insulin: 5/27
Diet alone: 4/11
RR = 0.51 (95% CI 0.07 to 3.71)*

Diet alone vs. no diet
Diet alone: 4/11
No diet: 9/34
RR = 1.37 (95% CI 0.52 to 3.58)*

Shoulder dystocia
Diet + insulin vs. diet alone
Diet + insulin: 0/27
Diet alone: 0/11</td>
<td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). No - the first 20 participants were not allocatedly randomly.

A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). No

A3: The groups</td>
</tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Study dates July 1973 to February 1975. Source of funding Not reported.			Neonatal hypoglycaemia Macrosomia was arbitrarily defined as > 8.5lb (3.864kg) based on 15.2% of neonates of non-diabetic patients at the study centre being above this threshold. Neonatal hypoglycaemia was defined as < 30mg/100ml. Shoulder dystocia was not defined. Statistical analysis Not reported.	RR not calculable. Diet alone vs. no diet Diet alone: 0/11 No diet: 1/34 RR = 0.97 (95% CI 0.04 to 22.25)* Perinatal mortality Diet + insulin vs. diet alone Diet + insulin: 0/27 Diet alone: 0/11 RR not calculable. Diet alone vs. no diet Diet alone: 0/11 No diet: 0/34 RR not calculable. Hypoglycaemia Diet alone vs. no diet Diet alone: 0/11 No diet: 2/34 RR = 0.58 (95% CI 0.03 to 11.25)* *Calculated by the NCC-WCH technical team.	were comparable at baseline, including all major confounding and prognostic factors. Unclear - age is not reported. P-values are not quoted. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). No

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					C2: a. How many participants did not complete treatment in each group? None. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Yes C3: a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. No - shoulder dystocia was not defined. D3: A valid and reliable method was used to determine the outcome. Unclear D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Other information
None.</td></tr>
<tr><td>Full citation
Crowther,C.A., Hiller,J.E., Moss,J.R., McPhee,A.J., Jeffries,W.S., Robinson,J.S., Australian Carbohydrate Intolerance Study in Pregnant Women (ACHOIS) Trial Group., Effect of treatment of gestational diabetes mellitus on pregnancy outcomes, New England Journal of Medicine, 352, 2477-2486, 2005
Ref Id
66023
Country/ies where the study was carried out
Australia and the United Kingdom
Study type
Randomised controlled trial.</td><td>Sample size
The total sample size comprised 1000 women (490 intervention, 510 control).

Characteristics
<table><tr><td>Characteristic</td><td>Intervention</td><td>Control</td></tr><tr><td>Mean age, years</td><td>30.9 ± 5.4</td><td>30.1 ± 5.5</td></tr><tr><td>Body mass index*</td><td>26.8 (23.3 to 31.2)</td><td>26.0 (22.9 to 30.9)</td></tr><tr><td>Ethnicity, n (%)</td><td>White: 356 (73%)
Asian: 92 (19%)
Other: 42 (9%)</td><td>White: 396 (78%)
Asian: 72 (14%)
Other: 42 (8%)</td></tr><tr><td>Parous, n (%)</td><td>212 (43%)</td><td>251 (49%)</td></tr><tr><td></td><td></td><td></td></tr></table>
No p-values were reported.

*Body mass index reported as medians and IQRs.</td><td>Interventions
Intervention
Individualised dietary advice.
Instruction in self-monitoring of blood glucose (four times daily until within the recommended range for two weeks).
Insulin if required.

Recommended ranges for blood glucose:
Fasting glucose ≥ 3.5mmol/l (63mg/dl) and ≤ 5.5mmol/l (99mg/dl).
Pre-prandial glucose levels ≤ 5.5mmol/l (99mg/dl).
Two-hour post-prandial glucose levels ≤ 7.0mmol/l (126mg/dl).

Control
Clinical care as provided where</td><td>Details
18 collaborating centres (14 in Australia and 4 in the United Kingdom) participated in the study.

Eligible women were enrolled between 16 and 30 weeks' gestation.

Women were advised to follow a normal diet in the 48 hours before the oral glucose tolerance test (OGTT) and to fast in the preceding 8 hours.

Women assigned to the treatment group were informed that they had a diagnosis of glucose intolerance. Women assigned to the usual care group were informed that they did not have gestational diabetes. A proportion of the women who had a normal OGTT at screening were assigned to the usual care group to maintain blinding.

Women whose glucose</td><td>Results
Composite score: serious perinatal outcomes (n out of N total births)
Treatment: 7/506
Control: 23/524
Adjusted RR = 0.33 (95% CI 0.14 to 0.75)#

Shoulder dystocia (n out of N total births)
Treatment: 7/506
Control: 16/524
Adjusted RR = 0.46 (95% CI 0.19 to 1.10)#

Admission to neonatal nursery (n out of N total births)
Treatment: 357/506
Control: 321/524
Adjusted RR = 1.13 (95% CI 1.03 to 1.23)#

Large for gestational age (n out of N total births)
Treatment: 68/506
Control: 115/524
Adjusted RR = 0.62 (95% CI 0.47 to 0.71)#

Perinatal mortality
Treatment: 0/506</td><td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Unclear

A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear

A3: The groups were comparable at baseline, including</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Aim of the study To assess whether treatment of gestational diabetes reduces perinatal complications and/or affects maternal outcomes, mood or quality of life. Study dates September 1993 to June 2003. Source of funding Funded by research grants from: Medical Researc h Council Australia The Queen Victoria Hospital Research Foundation, Adelaide Supported by the Department of Obstetrics and Gynaecology at	Inclusion criteria Women with a single or twin pregnancy between 16 and 30 weeks' gestation who attended antenatal clinics at one of the collaborating hospitals and had ≥ 1 risk factor for GDM at screening or a positive 50g oral glucose challenge test and a 75g oral glucose tolerance test at 24 to 34 weeks' gestation. Cut-offs for the glucose tests were as follows: 50g oral glucose challenge test: glucose level one hour after challenge ≥ 7.8mmol/l (140mg/l). 75g oral glucose tolerance test: venous plasma glucose < 7.8mmol/l (140mg/dl) after an overnight fast and between 7.8 and 11.0mmol/l (198mg/dl) at two hours. Exclusion criteria Women with previously diagnosed GDM or active chronic systemic disease (except essential hypertension). Women with more a severe glucose impairment than the specified cut-offs for glucose tests.	screening for gestational diabetes is not available.	levels exceeded the pre-specified cut-offs were informed that they had gestational diabetes. Insulin was administered to women in the treatment group if: During the two week period where women monitored glucose two capillary fasting glucose results were ≥ 5.5mmol/l (99mg/dl), or At 35 weeks' gestation or less two post-prandial results were ≥ 7.0mmol/l (126mg/dl), or After 35 weeks' gestation post-prandial glucose was ≥ 8.0mmol/l (144mg/dl), or One capillary glucose result was ≥ 9.0mmol/l (162mg/dl) during the two week period Shoulder dystocia was assessed using a standardised checklist. Serious perinatal complications were defined as one or more of: death, shoulder dystocia, bone fracture or nerve palsy. Large for gestational age was defined as > 90th percentile. Hypoglycaemia levels	Control: 5/524 RR = 0.09 (95% CI 0.005 to 1.62)* Hypoglycaemia (n out of N total births) Treatment: 35/506 Control: 27/524 Adjusted RR = 1.42 (95% CI 0.87 to 2.32)# Treatment failure Treatment: 100/490 Control: 17/510 RR = 6.12 (95% CI 3.72 to 10.08)* Mode of delivery (n out of N women) Induction of labour Treatment: 189/490 Control: 150/510 Adjusted RR = 1.36 (95% CI 1.15 to 1.62)# Elective caesarean Treatment: 72/490 Control: 61/510 Adjusted RR = 1.17 (95% CI 0.85 to 1.60)# Emergency caesarean Treatment: 80/490 Control: 103/510 Adjusted RR = 0.87 (95% CI 0.68 to 1.13)#	all major confounding and prognostic factors. Unclear B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. No B2: Participants receiving care were kept 'blind' to treatment allocation. No - however the control group did not know their diagnosis. B3: Individuals administering care were kept 'blind' to treatment allocation. No - see point B2. C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
the University of Adelaide.			requiring therapy were determined by the attending physician. Perinatal death was not defined. Statistical analysis An intention to treat analysis was used. For binary outcomes adjusted relative risks and 95% confidence intervals were calculated using log binomial regression. Continuous variables were analysed using ANOVA if normally distributed or non-parametric tests where appropriate. No adjustment was made for clustering by mother for twin pregnancies as no evidence of increased variance was identified. P-values < 0.05 were considered significant. Sidak's adjustment was used for multiple end point analyses. A sample size of 1000 was calculated for 80%	#Results from log binomial regression adjusted for maternal age, ethnicity and parity. *Calculated by the NCC-WCH technical team.	C2: a. How many participants did not complete treatment in each group? None b. The groups were comparable for treatment completion. Yes C3: a. For how many participants in each group were no outcome data available? None b. The groups were comparable with respect to the availability of outcome data. Yes D. Detection bias D1: The study had an appropriate length of follow-up. Unclear D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			power at the 5% level to detect a reduction in the risk of a serious perinatal outcome from 5.2% to 2.0%, based on outcomes reported for all South Australian births. A pre-specified stopping rule was put in place for a difference in major end points of ≥ 3 SD between groups.		was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information None.
Full citation Cypryk,K., Kaminska,P., Kosinski,M., Pertynska-Marczewska,M., Lewinski,A., A comparison of the effectiveness, tolerability and safety of high and low carbohydrate diets in women with gestational	Sample size Total sample size comprised 30 women (15 intervention, 15 control). Characteristics All women were Caucasian. Inclusion criteria Diagnosis of gestational diabetes according to WHO criteria.	Interventions Intervention 45% of daily intake was from carbohydrates, 25% protein and 30% from fat. Control 60% of daily intake was from carbohydrates, 25% protein and 15% from fat.	Details Before allocation to the prescribed diets, glycaemic levels were obtained from patients' diaries from the previous 3 to 4 days. This aimed to obtain an average 24 hour glycaemia value under normal conditions. Participants were then randomised to either diet. All participants received	Results Caesarean delivery Low carbohydrate: 7/15 High carbohydrate: 5/15 RR = 1.40 (95% CI 0.57 to 3.43)* Vaginal delivery Low carbohydrate: 7/15 High carbohydrate: 9/15 RR = 0.77 (95% CI 0.39 to 1.52)* Macrosomia	Limitations NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
diabetes, Endokrynologia Polska, 58, 314-319, 2007 Ref Id 177190 Country/ies where the study was carried out Poland Study type Randomised controlled trial. Aim of the study To evaluate the effectiveness and safety of high and low carbohydrate diets in women with gestational diabetes mellitus. Study dates Not reported. Source of funding Not reported.	Exclusion criteria Not reported.		education from a dietician and agreed to follow the prescribed diets for 14 days during which time SMBG was undertaken four times per day (fasting and 2 hours after each main meal). After assessment of food diaries on day 15 participants were asked to continue the diet until delivery. Targets for glucose during pregnancy were ≤ 90mg/dl fasting and ≤ 120mg/dl 2 hours post-prandial. Statistical analysis Between group comparisons of glycaemia were made using independent Student's t-tests or Mann-Whitney U tests where appropriate. P-values < 0.05 were considered significant.	Low carbohydrate: 0/15 High carbohydrate: 0/15 RR not calculable. *Calculated by the NCC-WCH technical team.	(which would have balanced any confounding factors equally across groups). Unclear A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear - no demographic data were provided. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					allocation. Unclear B3: Individuals administering care were kept 'blind' to treatment allocation. Unclear C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? None. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Yes C3:

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. No - "physiological" delivery was not defined. D3: A valid and reliable method was used to determine the

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					outcome. Unclear D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information None.
Full citation de Barros,M.C., Lopes,M.A., Francisco,R.P., Sapienza,A.D., Zugaib,M., Resistance exercise and glycemic control in women with gestational diabetes mellitus, American Journal of Obstetrics & Gynecology, 203, 556-556, 2010	Sample size Total sample size comprised 64 women (32 intervention, 32 control). Characteristics Not reported. Inclusion criteria Not reported. Exclusion criteria Not reported.	Interventions Intervention Participants performed resistance exercise using a resistance band. Exercise comprised a series of eight circuit-based activities. Women performed 15 reps of each exercise with a maximum of one minute's rest between each exercise. Women	Details Women were randomised into either treatment group. Participants in the intervention group received written instructions in how to perform each exercise. Glycaemic profiles of all participants were determined weekly. Insulin therapy was initiated when more than 30% of glucose	Results Requirement for insulin therapy Intervention: 7/32 Control: 18/32 RR = 0.38 (95% CI 0.18 to 0.78)* *Calculated by the NCC-WCH technical team.	Limitations NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Ref Id 145076 Country/ies where the study was carried out Brazil Study type Randomised controlled trial. Aim of the study To assess the impact of resistance exersise on insulin requirements in women with gestational diabetes mellitus. Study dates October 2006 to November 2008. Source of funding Not reported.		progressed from 2 circuits initially to 3 circuits after 3 weeks of inclusion. Control No exercise programme.	measurements were above the recommended value or when 20 to 30% of measurements were above the recommended value and foetal weight was > 75th percentile. Diagnosis criteria for GDM were not defined. Statistical analysis X2 tests were used to analyse categorical variables, Student's t-tests were used to analyse continuous variables.		confounding factors equally across groups). Unclear - randomisation method was not described. A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear - no baseline characteristics were reported. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Unclear B2: Participants receiving care were

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					kept 'blind' to treatment allocation. Unclear B3: Individuals administering care were kept 'blind' to treatment allocation. Unclear C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? Not reported. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Unclear

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					C3: a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. No - criteria for initiating insulin therapy were not reported. D3: A valid and reliable method

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					was used to determine the outcome. Unclear D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information None.
Full citation Garner,P., Okun,N., Keely,E., Wells,G., Perkins,S., Sylvain,J., Belcher,J., A randomized controlled trial of strict glycemic control and tertiary level obstetric care versus routine obstetric care in	Sample size The total sample size comprised 300 women (150 intervention, 150 control). Characteristics Characteristic / Treatment / Control / P-value Mean pre-pregnancy weight, kg / 68.91 ± 16.87 / 71.23 ± 19.78 / 0.28 Mean age, / 30.7 ± 4.8 / 30.7 ± / 0.98	Interventions Intervention Standard obstetric care and strict glycaemic control: Counselling 35kcal/kg/day intake Instruction in self monitoring of blood glucose Control Standard obstetric	Details The study was undertaken at two teaching hospitals in Ottawa. The goals of the pilot study were to assess patient acceptance, determine realistic enrollment rates, streamline data collection and identify adverse events in the standard care group. Women randomised to	Results Macrosomia Treatment: 6/149 Control: 6/150 RR = 1.01 (95% CI 0.33 to 3.06)* Neonatal hypoglycaemia Treatment: 21/149 Control: 13/150 RR = 1.73 (95% CI 0.91 to 3.30)* Vaginal delivery	Limitations NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
the management of gestational diabetes: a pilot study, American Journal of Obstetrics and Gynecology, 177, 190-195, 1997 Ref Id 153220 Country/ies where the study was carried out Canada Study type Randomised controlled trial pilot study. Aim of the study To undertake a pilot study in order to determine whether intensive obstetric-medical treatment reduced the risk of foetal macrosomia in women with gestational diabetes mellitus compared with routine obstetric care.	years 4.6 Inclusion criteria All pregnant women diagnosed with gestational diabetes between 24 and 32 weeks' gestation. Diagnosis of GDM was made using a 75g glucose screening test between 24 and 28 weeks' gestation with a one hour cut-off of 8.0mmol/l (145mg/dl). Women with a positive screening test undertook an oral glucose tolerance test using 75g glucose. All women diagnosed with GDM were assessed at a clinic and eligible women were enrolled. Exclusion criteria Multiple gestation, maternal-foetal blood group incompatibility, known congenital abnormality, prior evidence of placenta previa or abruptio placentae, significant maternal disease (including chronic hypertension, connective tissue disease, endocrine disorders and chronic hepatic disease), long-term medical therapy affecting glucose metabolism and imminent delivery.	care.	the treatment group were followed up in tertiary care bi-weekly. Targets for blood glucose were fasting levels < 4.4mmol/l (80mg/dl) and one hour post-prandial levels < 7.8mmol/l (140mg/dl). Targets were achieved in all women. If values were exceeded on two or more occasions insulin therapy was initiated. Women randomised to the control group were asked to continue a normal healthy diet for pregnancy as recommended by the Canada Food Guide. Two glucose tests per week were taken for comparison with the treatment group. Results were telephoned to an independent observer. Patients returned to their normal obstetric care provider. A "failed" control group of women with previously undiagnosed type 1 or type 2 diabetes was identified. It was considered unethical not to treat these women therefore they were transferred to the treatment arm if fasting	Treatment: 118/149 Control: 121/150 RR = 0.91 (95% CI 0.81 to 1.02)* Caesarean delivery Treatment: 30/149 Control: 28/150 RR = 1.10 (95% CI 0.69 to 1.75)* Perinatal mortality Treatment: 0/149 Control: 0/150 RR not calculable. Treatment failure Treatment: 36/149 Control: not reported RR = not calculable *Calculated by the NCC-WCH technical team.	(which would have balanced any confounding factors equally across groups). Unclear A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear - parity and ethnicity not reported. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. No B2: Participants receiving care were kept 'blind' to treatment

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Study dates September 1991 to May 1994. Source of funding Not reported.			capillary glucose levels were > 7.8mmol/l (140mg/dl) or one hour post-prandial levels were > 11.1mmol/l (200mg/dl). Foetal macrosomia was defined as > 4500g, regardless of gestational age. Perinatal mortality and neonatal hypoglycaemia were not defined. Statistical analysis Data were analysed using the intention to treat principle. For discrete outcomes data were summarised using percentages and groups were compared using X2 or Fisher's exact tests. Means of continuous outcomes were compared between groups using Student's t-tests or the Wilcoxon sign rank test. The sample size of 300 was not sufficient to detect statistically significant differences between treatment groups for macrosomia		allocation. Unclear B3: Individuals administering care were kept 'blind' to treatment allocation. Unclear C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? 1 lost to follow-up in the intervention group, 0 in the control group. b. The groups were comparable for treatment completion. Yes C3: a. For how many participants in each group were no outcome data available? 1 lost to

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			rates, operative deliveries or adverse foetal or neonatal outcomes.		follow-up in the intervention group, 0 in the control group. b. The groups were comparable with respect to the availability of outcome data. Yes D. Detection bias D1: The study had an appropriate length of follow-up. Unclear D2: The study used a precise definition of outcome. No - definitions not provided for all outcomes. D3: A valid and reliable method was used to determine the outcome. Unclear D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>other important confounding and prognostic factors. Unclear

Other information
None.</td></tr>
<tr><td>Full citation
Grant,S.M., Wolever,T.M., O'Connor,D.L., Nisenbaum,R., Josse,R.G., Effect of a low glycaemic index diet on blood glucose in women with gestational hyperglycaemia, Diabetes Research and Clinical Practice, 91, 15-22, 2011
Ref Id
157375
Country/ies where the study was carried out
Canada
Study type
Randomised pilot study

Aim of the study</td><td>Sample size
N = 43

Characteristics
<table><tr><td>Characteristic</td><td>Control</td><td>Low GI</td><td>P-value</td></tr><tr><td>Diagnosis (GDM:IGTP)</td><td>17:6</td><td>15:9</td><td>NS</td></tr><tr><td>Non-Caucasian ethnicity, n (%)</td><td>19 (82.6%)</td><td>21 (83.3%)</td><td>NS</td></tr><tr><td>Mean maternal age, years</td><td>34 ± 1.1</td><td>34 ± 0.1</td><td>NS</td></tr><tr><td>Mean gestational age at diagnosis, weeks</td><td>27 ± 0.5</td><td>27 ± 0.7</td><td>NS</td></tr><tr><td>Mean pre-pregnancy BMI, kg/m2</td><td>26 ± 1</td><td>27 ± 1</td><td>NS</td></tr><tr><td>Mean HbA_{1c}, %</td><td>5.4 ± 0.1</td><td>5.3 ± 0.1</td><td>NS</td></tr></table>
Data presented as mean ± SE.</td><td>Interventions
Intervention
Standard nutrition therapy for women with gestational hyperglycaemia with low GI starch content.

Control
Standard nutrition therapy for women with gestational hyperglycaemia with intermediate to high GI starch content.</td><td>Details
The study was a randomised, open-label pilot which aimed to recruit a total of 50 women. Women were stratified according to whether they were diagnosed with GDM or impaired glucose tolerance of pregnancy (IGTP). A total of 47 women were randomised. Four women withdrew during the run-in period before treatments commenced.

Standard therapy comprised patients being introduced to the Diabetes Food Guide and Canadian dietary recommendations for a healthy pregnancy. Starch choices and servings were recommended to each woman by the clinic</td><td>Results
Large for gestational age, n/N
Low GI: 2/18
Control: 3/20
RR = 0.74 (95% CI 0.13 to 4.18)*

Treatment with insulin, n/N
Low GI: 13/18
Control: 12/20
RR = 1.20 (95% CI 0.75 to 1.93)*

*Calculated by the NCC-WCH technical team.</td><td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Unclear

A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
To evaluate the effects of a low glycaemic index diet in women with gestational hyperglycaemia. This pilot study aimed to test the feasibility of the intervention and determine its effect on fasting serum glucose, HbA_{1c} and SMBG and obtain preliminary data on infant birth weight. Study dates April 2006 to January 2007 Source of funding Supported by the Danone Institute of Canada.	Exact p-values were not reported for non-significant results. Overall ethnicity, % South East Asian = 25% Indian = 21% Caucasian = 21% East Asian = 11% Caribbean = 9% Hispanic = 6% Mixed = 6% Inclusion criteria Aged 18 to 45 years Diagnosed with gestational hyperglycaemia (GDM or impaired glucose tolerance of pregnancy) Referred to the Diabetes in Pregnancy Clinic at St Michael's Hospital Willing and able to comply with the study protocol and to provide written consent Exclusion criteria Multiple pregnancies An acute or chronic illness affecting carbohydrate metabolism Presence of type 1 or type 2 diabetes prior to the current pregnancy Use of insulin prior to providing consent > 34 weeks' gestation Unable to communicate in English with no translator available		dietician. Women in the study were asked to select their starch choices from a specific exchange list depending upon their treatment group allocation. The control group received a choice of intermediate and high GI foods reflecting the usual intake of a woman with gestational hyperglycaemia. Women in the low GI group chose from a list of foods with low glycaemic index. Women were not advised about food types other than starchy foods. Primary outcome measures were fasting serum glucose and HbA_{1c} assessed at baseline and at 4 weeks and SMBG from baseline to week 8. Blood glucose was measured four times daily by women (fasting and 2 hours after breakfast, lunch and dinner). If targets for SMBG were not met using either the intervention or control treatments insulin was prescribed. The decision to administer insulin was		or treatment allocation). Yes A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			made by a clinician blinded to allocation. The target range for blood glucose was that recommended by the Canadian Diabetes Association: Fasting 3.8 to 5.2mmol/l 2 hour postprandial 5.0 to 6.6mmol/l Women were followed from recruitment to delivery. Five women dropped out during the treatment period leaving a total of 38 women with data on birth weight. Large for gestational age was defined as > 90th percentile for sex and gestational age. Statistical analysis Data were analysed on an intention-to-treat basis. P-values < 0.05 were taken to be statistically significant.		length of follow-up). Yes C2: a. How many participants did not complete treatment in each group? Six in the low GI group and three in the control group did not complete treatment. Five of these women dropped out after randomisation but the distribution between treatment groups was not reported. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Unclear C3: a. For how many participants in each group were no outcome data available?

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					Analyses are based on women with available data only. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Yes D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>intervention.
Unclear

D5: Investigators were kept 'blind' to other important confounding and prognostic factors.
Unclear

Other information
Pilot study therefore underpowered to detect associations.</td></tr>
<tr><td>Full citation
Hague,W.M., Davoren,P.M., Oliver,J., Rowan,J., Contraindications to use of metformin. Metformin may be useful in gestational diabetes, BMJ (Clinical research ed.), 326, 762-, 2003
Ref Id
177294
Country/ies where the study was carried out
Australia</td><td>Sample size
n=30

Characteristics
Women were matched for age, parity BMI and gestational age at entry to the study.
<table>
<tr><td>Characteristic</td><td>Metformin (n=16)</td><td>Insulin (n=14)</td></tr>
<tr><td>Maternal age (years)</td><td>33.7 (4.44)</td><td>34.1 (3.70)</td></tr>
<tr><td>Median parity (range)</td><td>1 (0-4)</td><td>1 (0-5)</td></tr>
<tr><td>Maternal BMI at trial entry</td><td>39.5 (6.94)</td><td>37.9 (6.87)</td></tr>
<tr><td>Gestation at time of diagnosis</td><td>25.8 (5.51)</td><td>27.6 (3.80)</td></tr>
</table></td><td>Interventions
Metformin and insulin were the treatments compared but no further details of these treatments are given. No details of any concurrent dietary interventions or monitoring techniques are presented.</td><td>Details
Not stated

Statistical analysis
Between-group differences in mean C-peptide levels were compared using Mann-Whitney U tests. No other statistical methods were reported.</td><td>Results
<table>
<tr><td>Outcome</td><td>Metformin (n=16)</td><td>Insulin (n=14)</td></tr>
<tr><td>Vaginal delivery (%)</td><td>5 (31%)</td><td>11(79%)</td></tr>
<tr><td>Induction of labour (%)</td><td>5 (31%)</td><td>9 (64%)</td></tr>
<tr><td>Elective Caesarean section (%)</td><td>8 (50%)</td><td>2 (145)</td></tr>
<tr><td>Emergency Caesarean section (%)</td><td>2 (13%)</td><td>1 (7%)</td></tr>
</table></td><td>Limitations
NICE guidelines manual Appendix C
Methodology checklist: randomised controlled trials
Appropriate randomisation method: unclear, not stated
Adequate allocation concealment: unclear, not stated
Groups comparable at baseline: unclear, not stated
Groups received the same care (apart from the intervention): unclear, not stated</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Study type Pilot randomised controlled trial Aim of the study To compare the effects of insulin and metformin on outcomes in a population of women with gestational diabetes Study dates Not stated Source of funding Not stated	<table><tr><td>OGTT Fasting blood glucose</td><td>5.6 (1.26)</td><td>5.4 (0.52)</td></tr><tr><td>OGTT 2h post load glucose</td><td>10.0 (2.07)</td><td>9.4 (1.42)</td></tr></table> Inclusion criteria Women diagnosed with gestational diabetes according to ADIPS criteria and who gave consent to participate Exclusion criteria Not stated			<table><tr><td>Birth weight >4000g</td><td>2</td><td>2</td></tr><tr><td>Neonates requiring IV dextrose</td><td>4</td><td>1</td></tr></table>	Participants kept 'blind' to allocation: no, not possible Care givers kept 'blind' to allocation: no, not possible Follow up equal for groups: yes How many participants did not complete treatment in each group?: none Were the groups were comparable for treatment completion: yes For how many participants in each group were no outcome data available?: none The groups were comparable with respect to the availability of outcome data: yes Appropriate length of follow-up: yes Precise outcome definitions used: unclear, no definitions provided Outcome determined using valid and reliable methods: yes Investigators kept 'blind' to allocation: unclear, not stated

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					Investigators kept 'blind' to other important confounding and prognostic factors: unclear, not stated Other information None.
Full citation Ijas,H., Vaarasmaki,M., Morin-Papunen,L., Keravuo,R., Ebeling,T., Saarela,T., Raudaskoski,T., Metformin should be considered in the treatment of gestational diabetes: a prospective randomised study, BJOG: An International Journal of Obstetrics and Gynaecology, 118, 880-885, 2011 Ref Id 155747 Country/ies where the study was carried out	Sample size Of 239 women referred to outpatient clinics in the 2 study hospitals, 128 women were eligible for inclusion and 100 agreed to participate. Sample size calculation is presented: to detect a 30% unit difference in macrosomia rates between the study groups, a two sided test with 80% power and significance level of 0.05, a sample size of 50 women in each group was needed. Characteristics Metformin group n=47 Insulin group n=50 Age (years) Metformin group = 32.3 ± 5.6 Insulin group = 31.7 ± 6.1 Parity Metformin group = 1.6 ± 2.4 Insulin group = 1.6 ±1.8 Nulliparous Metformin group = 18 (38.2%) Insulin group = 16 (32%) BMI at first antenatal visit Metformin group = 31.5 ± 6.5	Interventions Women were randomised to treatment with metformin (n=50) or insulin (n=50) following tests to ensure normal renal and liver functioning. Metformin was started at 750mg once/day in the first week, 750mg twice/day in the second week and 750mg three times/day from the third week onwards. Medication was discontinued if significant side effects (eg diarrhoea) occurred. Supplemental insulin was added	Details All women received dietary and lifestyle counselling. Home monitoring of glucose concentrations were performed twice weekly using 4-6 point daily profiles. Target concentrations were <5.3 mmol/l for fasting and <6.7 mmol/l for postprandial glucose. Glucose concentrations were reported to the diabetes nurse at 2 to 4 week intervals. If fasting or postprandial concentrations exceeded target levels at least twice , then pharmacological treatment was considered. Participants were followed at outpatient clinics every 4 weeks (gsetational age 12-32	Results Spontaneous vaginal delivery Metformin group = 22/47 (46.8%) Insulin group = 36/50 (72%) RR = 0.8 (95% CI 0.46 to 0.92) p=0.011 Labour induction Metformin group = 24/47 (51.0%) Insulin group = 26/50 (52%) RR = 1.0 (95% CI 0.67 - 1.45) p= 0.960 Vacuum extraction Metformin group = 7/47 (14.9%) Insulin group = 4/50 (8%) p=0.041 Caesarean section Metformin group = 18/47 (38.3%) Insulin group = 10/50 (20%) RR = 1.9 (95% CI 0.99 to 3.31) p=0.047 Need for additional insulin	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials Appropriate randomisation method: Manually generated randomisation code Adequate allocation concealment: Yes (opaque envelopes) Groups comparable at baseline: Yes Groups received the same care (apart from the intervention): Yes Participants kept 'blind' to allocation: No, not possible Care givers kept 'blind' to allocation: No, not possible Follow up equal for

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Finland Study type Randomised controlled trial Aim of the study To investgate whether metformin is as effective as insulin in preventing foetal macrosomia in women with gestational diabetes Study dates 22 June 2005 to 30 June 2009 Source of funding The Foundation of Alma and KA Snellman, Oulu, Finland	Insulin group = 30.8 ± 5.4 Fasting glucose in OGTT (mmol/l) Metformin group = 5.6 ± 0.9 Insulin group = 5.4 ± 0.6 2 hour glucose in OGTT (mmol/l) Metformin group = 8.2 ± 1.9 Insulin group = 8.1 ± 1.8 Gestational age at OGTT (weeks) Metformin group = 23 ± 5.7 Insulin group = 23 ± 5.4 Gestational age at randomisation (weeks) Metformin group = 30 ± 4.9 Insulin group = 30 ± 4.0 HbA_{1c} at randomisation (weeks) Metformin group = 5.9 ± 0.4 Insulin group = 5.9 ±0.4 There were no significant differences in any baseline characteristics between the two groups Inclusion criteria All women with risk factors for gestational diabetes underwent a 75g OGTT as part of free primary health care. Women were tested if they had any of the following: body mass index over 25 kg/m2, aged over 40 years, a previous baby over 4500g, glucosuria during pregnancy, previous gestational diabetes, or suspected foetal macrosomia in current pregnancy. Women who tested positive were referred to outpatient maternity clinics. Women who were diagnosed with gestational diabetes between 12 and 34 weeks of gestation and with singleton pregnancies were included in the study. Exclusion criteria The presence of pre-eclampsia, essential hypertension requiring antihypertensive	if normoglycaemia was not achieved in the 1-2 weeks using the maximum dose. Insulin treatment consisted of long acting insulin to normalise fasting glucose concentrations and rapid acting insulin to normalise postprandial glucose concentrations. Women continued to measure daily profiles of capillary glucose concentrations twice a week and reported values to the diabetes nurse.	weeks), every 2 weeks (gestational age 32-36 weeks) or once or twice weekly (after gestational age 36 weeks). At every visit, maternal weight gain was recorded and foetal growth was investigated using ultrasound. HbA_{1c} was measures at randomisation, 2 weeks after initiation of treatment and monthly thereafter. Statistical analysis Sample size calculations were designed to detect a 30% difference in macrosomia rates. Based on 80% power and a significance level of 0.05 the required sample size was 50 women per arm. Between-group comparisons were made using Student's t-tests or Mann-Whitney U tests for continuous data. Fisher's exact tests or X2 tests were used to analyse categorical data. Analyses were two-tailed and p-values < 0.05 were considered to be significant.	Metformin group = 15/47 (31.9%) required supplemental insulin to reach normoglycaemia. 3/15 women discontinued metformin because of gastrointestinal side effects and changed to insulin. 1/47 changed to insulin after 3 weeks because of elevated liver enzymes. 1/47 had a reduced dose of metforming sue to side effects (diarrhoea). Both these women were analysed in the metformin group Large for gestational age infants (Definition: birthweight greter then +2SDs using Finnish specific charts adjusted for gestational age) Metformin group = 4/47 (8.5%) Insulin group = 5/50 (10%) RR = 0.9 (95% CI 0.24 to 2.98) p= 0.901 Neonates transferred to NICU Metformin group = 7/47 (14.9%) Insulin group = 11/50 (22%) RR = 0.7 (95% CI 0.29 to 1.60) p= 0.368 Neonatal hypoglycaemia (Definition: hypoglycaemia that requires intravenous glucose treatment) Metformin group = 4/47 (8.5%)	groups: Yes How many participants did not complete treatment in each group?: Metformin 3/50, Insulin 0/50 Were the groups were comparable for treatment completion: yes For how many participants in each group were no outcome data available?: None The groups were comparable with respect to the availability of outcome data: Yes Appropriate length of follow-up: Yes Precise outcome definitions used: Yes Outcome determined using valid and reliable methods: Yes Investigators kept 'blind' to allocation: unclear, not stated Investigators kept 'blind' to other important confounding and prognostic factors:unclear, not stated

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	treatement and foetal growth restriction were criteria for exclusion from the study			Insulin group = 7/50 (14%) RR = 0.7 (95% CI 0.23 to 1.89) p=0.439 Birth injury (Definition: Clavicular fracture or brachial nerve injury) Metformin group = 0/47 Insulin group = (2/50 - both clavicular injuries following shoulder dysctocia) Perinatal mortality Metformin group = 0/47 Insulin group = 0/50	Other information None.
Full citation Lain,K.Y., Garabedian,M.J., Daftary,A., Jeyabalan,A., Neonatal adiposity following maternal treatment of gestational diabetes with glyburide compared with insulin, American Journal of Obstetrics & Gynecology, 200, 501-506, 2009 Ref Id 144548 Country/ies where the study	Sample size 99 women were randomised (Glibenclamide n=49, Insulin N=50) and results for neonatal measure of growth (primary outcomes) are presented for 82 babies, 41 in each group. No details regarding the women lost to follow up are provided Characteristics The groups had similar baseline characteristics at entry to the study including gestational age at randomisation, 3 hour OGTT results and baseline HbA_{1c} Inclusion criteria Pregnant women who had abnormal results from a screen using a 50g 1 hour glucose challenge test (135mg/dl) and who went on to have a 3 hour OGTT. Women who had two abnormal values, an elevated fasting value from the 3 hour OGTT or those with a 1 hour post glucose load OGTT value of >200mg/dl were diagnosed with gestational diabetes and included in the study.	Interventions No details of diet, exercise or monitoring techniques are presented Glibenclamide doses started at 2.5mg/day and were increased by 2.5-5mg weekly. Doses were taken once or twice daily. If a maximum dose of 20mg/day glibenclamide did not achieve goals, then women were transitioned to insulin. Insulin doses started at 0.8U/kg	Details No details of randomisation are presented. Neonatal measuements were performed in triplicate within the first 36 hours of life. Infant birthweights were compared with insitutionally derived standards stratified by race and sex. Statistical analysis Not reported.	Results Treatment failure Glibenclamide = 3/49 women who were transitioned to insulin Large for gestational age Glibenclamide = 12/41 Insulin = 3/38 Admission to NICU Glibenclamide = 6/49 Insulin = 5/50 Neonatal hypoglycaemia Glibenclamide = 4/49 Insulin = 0/50 Shoulder dystocia Glibenclamide = 1/49 Insulin = 2/50 Intrauterine death Glibenclamide = 1/40 (associated with trisomy 21) Insulin = 0/50 Neonatal death Glibenclamide = 0/49 Insulin = 0/50	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials Appropriate randomisation method: unclear, not stated Adequate allocation concealment: unclear, not stated Groups comparable at baseline: yes Groups received the same care (apart from the intervention): yes Participants kept 'blind' to allocation: no Care givers kept 'blind' to allocation:

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
was carried out United States of America Study type Randomised controlled trial Aim of the study To examine neonatal body composition and metabolic markers at birth in women with gestational diabetes who were treated with glibenclamide or insulin Study dates 2002 to 2005 Source of funding Grants from the American Association of Obstetricians and Gynaecologists Foundation and the Magee Womens Health Foundation	Exclusion criteria Not presented	administered in multiple daily injections and were increased up to twice weekly as necessary. Women receiving glibenclamide were transitioned to insulin if the maximum dise of 20mg/day did not achieve targets.			no Follow up equal for groups: yes How many participants did not complete treatment in each group?: none Were the groups comparable for treatment completion: yes For how many participants in each group were no outcome data available?: Depending on outcome, up to 13 were lost from the insulin group and up to 8 in the glibenclamide group The groups were comparable with respect to the availability of outcome data: yes Appropriate length of follow-up: yes Precise outcome definitions used: no, precise definitions are not presented for all outcomes, especially shoulder dystocia Outcome determined using

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>valid and reliable methods: unclear
Investigators kept 'blind' to allocation: unclear, not stated
Investigators kept 'blind' to other important confounding and prognostic factors: unclear, not stated

Other information
None.</td></tr>
<tr><td>Full citation
Landon,M.B., Spong,C.Y., Thom,E., Carpenter,M.W., Ramin,S.M., Casey,B., Wapner,R.J., Varner,M.W., Rouse,D.J., Thorp,J.M.,Jr., Sciscione,A., Catalano,P., Harper,M., Saade,G., Lain,K.Y., Sorokin,Y., Peaceman,A.M., Tolosa,J.E., Anderson,G.B., Eunice Kennedy Shriver National Institute of Child Health and</td><td>Sample size
The total sample size comprised 958 women (485 intervention group, 473 control).

Characteristics
<table><tr><td>Characteristic</td><td>Treatment</td><td>Control</td></tr><tr><td>Mean age, years</td><td>29.2 ± 5.7</td><td>28.9 ± 5.6</td></tr><tr><td>Primigravida, n (%)</td><td>104 (21.4%)</td><td>123 (26.0%)</td></tr><tr><td>Race/ethnic group, n (%)</td><td></td><td></td></tr><tr><td>Black</td><td>56 (11.5%)</td><td>54 (11.4%)</td></tr><tr><td>White</td><td>123 (25.4%)</td><td>119 (25.2%)</td></tr><tr><td>Asian</td><td>22 (4.5%)</td><td>28</td></tr></table></td><td>Interventions
Intervention
Dietary counselling and therapy.
Instruction in self monitoring of blood glucse.
Insulin where appropriate.

Control
Standard obstetric care.</td><td>Details
After an overnight fast eligible women completed a blinded 3 hour 100g oral glucose tolerance test.

Women who met these criteria were randomly assigned to each group using minimisation, stratified by clinical centre. Out of 19665 who had abnormal glucose loading tests, 10989 met inclusion criteria and 7381 consented to an OGTT. Of these women 1889 were enrolled into the trial. This included a cohort of women who had positive 50g glucose</td><td>Results
Composite outcome: hypoglycaemia, hyperbilrubinaemia, elevated cord blood C-peptide, stillbirth/neonatal death, birth trauma
Treatment: 149/460
Control: 163/440
RR = 0.87 (95% CI 0.72 to 1.07)

Hyperinsulinaemia
Treatment: 75/423
Control: 92/403
RR = 0.78 (97% CI 0.57 to 1.05)

Large for gestational age
Treatment: 34/477
Control: 66/454</td><td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). No - minimisation was used.</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Human Development Maternal-Fetal Medicine Units Network., A multicenter, randomized trial of treatment for mild gestational diabetes, New England Journal of Medicine, 361, 1339-1348, 2009 Ref Id 155651 Country/ies where the study was carried out United States of America Study type Randomised controlled trial. Aim of the study To determine whether treatment of women with mild gestational diabetes reduces perinatal and obstetric complications. Study dates October 2002 to	<table><tr><td></td><td></td><td>(5.9%)</td></tr><tr><td>Hispanic</td><td>281 (57.9%)</td><td>265 (56.0%)</td></tr><tr><td>Other</td><td>3 (0.6%)</td><td>7 (1.5%)</td></tr><tr><td>Body mass index at baseline</td><td>30.1 ± 5.0</td><td>30.2 ± 5.1</td></tr></table> No p-values were reported. Inclusion criteria Women were included if, between 24 weeks 0 days and 20 weeks 6 days' gestation they had a blood glucose between 7.5mmol/l (135mg/dl) and 11.1mmol/l (200mg/dl) one hour after a 50g oral glucose loading (screening) test. Mild GDM was defined as a fasting glucose < 5.3mmol/l and two or three timed measurements that exceeded the following thresholds: One hour > 10.0mmol/l (180mg/dl) Two hours > 8.6mmol/l (155mg/dl) Three hours > 7.8mmol/l (140mg/dl) Exclusion criteria Pre-existing diabetes, an abnormal result on a glucose screening test before 24 weeks' gestation, prior gestational diabetes, a history of stillbirth, multifoetal gestation, asthma or chronic hypertension, corticosteroid treatment or if imminent/pre-term delivery was likely because of maternal or foetal conditions. To capture only mild GDM women with an OGTT		loading tests but a normal oral glucose tolerance test were matched with the study cohort according to BMI and race and included in the control group in order to maintain blinding (n = 931). Insulin was prescribed if the majority of fasting or post-prandial values were > 5.3mmol/l (95mg/dl) or > 6.7mmol/l (120mg/dl), respectively. The primary study outcome was a composite outcome which included: Perinatal mortality (stillbirth or neonatal death) Hypoglycaemia Hyperbilirubinaemia Neonatal hyperinsulinaemia Birth trauma Hyperinsulinaemia was defined as cord-blood C-peptide > 95th percentile. Neonatal hypoglycaemia was defined as glucose < 1.9mmol/l (35mg/dl) two hours after birth and before	RR = 0.49 (97% CI 0.32 to 0.76) Induction of labour Treatment: 130/476 Control: 122/455 RR = 1.02 (97% CI 0.81 to 1.29) Caesarean delivery Treatment: 128/476 Control: 154/455 RR = 0.79 (97% CI 0.64 to 0.99) Shoulder dystocia Treatment: 7/476 Control: 18/455 RR = 0.37 (97% CI 0.14 to 0.97) Perinatal mortality Treatment: 0/485 Control: 0/473 RR not calculable. Treatment failure Treatment: 37/476 Control: 2/455 RR = 17.68 (95% CI 4.29 to 72.93)* *Calculated by the NCC-WCH technical team.	A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear - no p-values reported. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Unclear B2: Participants receiving care were kept 'blind' to treatment allocation. Unclear B3: Individuals administering care were kept 'blind' to treatment

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
November 2007. Source of funding Supported by grants from the Eunice Kennedy Shriver National Institute of Child Health and Human Development., the General Clinical Research Centers and the National Center for Research Resources.	> 5.3mmol/l (95mg/dl) were excluded and their care provider informed.		feeding. Hyperbilirubinaemia was defined as serum bilirubin > 95th percentile. Birth trauma was defined as brachial plexus palsy or clavicular, humeral or skull fracture. Secondary neonatal outcomes: Birth weight > 4000g (macrosomia) Large for gestational age (> 90th percentile) Admission to the neonatal care unit Secondary maternal outcomes: Caesarean delivery Labour induction Shoulder dystocia (defined clinically) Statistical analysis Based on a literature review it was assumed that outcome rates would be between 20 and 30% in the control group. A composite outcome rate of 25% was assumed in the control group. Sample size was calculated to be 950 for a power of 80% to detect a 30% difference in the composite outcome with		allocation. No - blinded to diagnosis status of controls only. C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? None b. The groups were comparable for treatment completion. Yes C3: a. For how many participants in each group were no outcome data available? Unclear - missing data but numbers and/or group not reported. b. The groups were comparable with

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			treatment. Type 1 error was set at 5%. This sample size provided 85% power to detect a 30% reduction in rates of large for gestational age births and births > 4000g. Analyses were carried out according to the intention to treat principle. Categorical variables were compared using X2 or Fisher's exact tests. Continuous variables were analysed using the Wilcoxon ranksum test. An external data monitoring committee was used for four interim analyses. Adjusted type 1 error was calculated using the Lan-DeMets generalisation of the O'Brien-Fleming boundary. In final analyses p-values < 0.032 were considered significant, providing 97% confidence intervals for relative risks.		respect to the availability of outcome data. Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Unclear D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. No D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information None.

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td>Full citation
Langer,O., Anyaegbunam,A., Brustman,L., Divon,M., Management of women with one abnormal oral glucose tolerance test value reduces adverse outcome in pregnancy, American Journal of Obstetrics and Gynecology, 161, 593-599, 1989
Ref Id
180257
Country/ies where the study was carried out
United States of America
Study type
Randomised controlled trial.

Aim of the study
To determine the glycaemic profile in treated and untreated women with one abnormal value, to ascertain the</td>
<td>Sample size
N = 272

Characteristics
<table>
<tr><th>Characteristic</th><th>Treated (n = 63)</th><th>Untreated (n = 63)</th></tr>
<tr><td>Mean maternal age, years</td><td>31 ± 5</td><td>28 ± 6</td></tr>
<tr><td>Nulliparous, n (%)</td><td>18 (29%)</td><td>20 (32%)</td></tr>
<tr><td>Race, n (%)</td><td>-</td><td>-</td></tr>
<tr><td>Black</td><td>19 (30%)</td><td>21 (33%)</td></tr>
<tr><td>Hispanic</td><td>21 (33%)</td><td>21 (33%)</td></tr>
<tr><td>White</td><td>23 (36%)</td><td>21 (33%)</td></tr>
<tr><td>Mean gestational age at diagnosis, weeks</td><td>31 ± 3</td><td>31 ± 3</td></tr>
<tr><td>Obesity, n (%)</td><td>24 (38%)</td><td>26 (41%)</td></tr>
</table>
P-values were not reported for all values. Where they were reported the comparison included non-diabetic controls.

Inclusion criteria
All pregnant women routinely screened for GDM with one abnormal OGTT value
Between 24 and 28 weeks' gestation

Exclusion criteria
Not reported.</td>
<td>Interventions
Intervention
Diet comprising 25kcal/kg for women with a pre-pregnancy BMI ≥ 27 or 30kcal/kg for a BMI < 27.

Control
Women were instructed to continue their normal eating patterns.</td>
<td>Details
All women at the study centre were routinely screened using a 50g GCT. If one hour postprandial glucose was ≥ 130mg/dl (7.2mmol/l) women underwent a three hour OGTT.

A total of 272 women were included in the study. The main study group comprised 126 women with one abnormal OGTT value. These women were randomised into treated or untreated arms. A control group was also established of women with screening blood glucose < 140mg/dl (7.8mmol/l) and a normal OGTT result (n = 146).

All women in the randomised arms self-monitored capillary blood glucose seven times daily.

Women in the treatment group were advised to adhere to a diet comprising 25kcal/kg for those women with a BMI</td>
<td>Results
Large for gestational age
Diet: 4/63
No diet: 15/63
RR = 0.27 (95% CI 0.09 to 0.78)*

Neonatal hypoglycaemia
Diet: 1/63
No diet: 8/63
RR = 0.13 (95% CI 0.02 to 1.01)*

NICU admission
Diet: 4/63
No diet: 7/63
RR = 0.57 (95% CI 0.17 to 1.87)*

*Calculated by the NCC-WCH technical team.</td>
<td>Limitations

NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Unclear

A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear

A3: The groups were comparable at baseline, including all major confounding and prognostic factors.</td>
</tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
relationship between maternal and neonatal outcome in treated and untreated women with one abnormal OGTT value and to compare pregnancy outcome in normal women and women with one abnormal OGTT value. Study dates Not reported. Source of funding Part funded by an educational grant from Miles Laboratories.			≥ 27 or 30kcal/kg for those with a BMI < 27. All women were treated to achieve glycaemic control of < 95mg/dl (5.3mmol/l). This is assumed to be for fasting values though this is information is not reported in the study. When this was not achieved with diet alone, insulin was administered. Insulin dose was calculated as 0.7U/kg during pregnancy and given as MDI, two thirds in the morning and one third in the evening in split doses of regular and intermediate insulin. Women in the untreated group were advised to continue their normal eating habits. Women in this group were required to monitor capillary blood glucose for a baseline period of four weeks. Large for gestational age was defined as ≥ 90th percentile. Neonatal hypoglycaemia was defined as < 35mg/dl (1.9mmol/l).		Yes B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes C2: a. How many participants did not complete treatment in each group? None

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			NICU admission was recorded when length of stay was > 24 hours. Statistical analysis Pregnancy outcomes were compared between treatment groups and with the control group of non-diabetic women. Categorical data were analysed using χ2 tests or Fisher's exact test. Continuous data were analysed using Student't t test. Pearson's correlation coefficient was calculated for the relationship between glycaemic control and neonatal birthweight (percentile).		b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Yes C3: a. For how many participants in each group were no outcome data available? None b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Yes D. Detection bias D1: The study had an appropriate

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					length of follow-up. Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information Data for control subjects were not included in analyses as they are not relevant to the review protocol.
Full citation Langer,O.,	Sample size N= 404 women with gestational diabetes	Interventions Glibenclamide :	Details Diet : All women received	Results Treatment failure	Limitations NICE guidelines

Study details	Participants			Interventions	Methods	Outcomes and Results	Comments
Conway,D.L., Berkus,M.D., Xenakis,E.M., Gonzales,O., A comparison of glyburide and insulin in women with gestational diabetes mellitus, New England Journal of Medicine, 343, 1134-1138, 2000 Ref Id 177424 Country/ies where the study was carried out USA Study type Randomised controlled trial Aim of the study To evaluate whether glibenclamide might be an alternative to insulin therapy in women with gestational diabetes Study dates Not stated Source of	attending maternal health clinics in San Antonio Texas Glibenclamide group = 201 Insulin group = 203 Characteristics Mean age (yr): Glibenclamide (n=201) 29±7; (Insulin (n=203) 30±6 BMI ≥27.3 before pregnancy n (%): 141 (70%); 132 (65%) Nulliparity n (%): 56 (28%); 59 (29%) Family history on diabetes n (%): 86 (43%); 91 (45%) Previous gestational diabetes n (%): 24 (12%); 22 (11%) Previous infant with macrosomia n (%): 36 (18%); 45 (22%) Mean gestation at entry: 24±7; 25±7 Mean gestation at delivery: 38.7±1.6; 38.5±2.1 Mean clinic visits attended: 11±5; 12±6 Mean clinic visits missed: 1.5±2.1; 1.2±2.2 Mean blood glucose measurements/da: 4±2; 4±2			An initial dose of 2.5mg in the morning was increased in the first week by 2.5mg and by 5mg weekly thereafter if necessary to a maximum dose of 20mg/day. Blood glucose was reviewed in clinic weekly. Insulin : Insulin was started at a dosage of 0.7 units of insulin/kg actual body weight given subcutaneously, injected three times daily and increased as necessary to maintain targets. Treatment failure was defined taking the maximum dose without achieving glucose targets over a two week period. Oral medication was stopped in treatment failure and insulin therapy started.	dietary instruction for 3 meals and 4 snacks daily. Adherence was evaluated and reinforced at weekly clinic visits. The diet was designed to provide 30kcal/kg body weight for women of normal weight. Women who were obese (BMI>30) received a diet designed to deliver 25kcal/kg body weight. The calories were split by source with 40% from carbohydrates Monitoring : All women were trained to use a portable glucose meter at home and tested their blood glucose x7/day: in the morning (fasting value), before and 2 hours after lunch and dinner, at bedtime.Targets were fasting 60-90mg/dl; preprandial 80-95 mg/dl; 2 hour postprandial <120mg/dl. Blood glucose was measured for comparison at weekly clinic. Statistical analysis An intention-to-treat analysis was performed. X2 tests were performed to compare categorical data between treatment groups and Student's t-	Glibenclamide group = 8/201 (4%) Large for gestational age (Birth weight >90th percentile) Glibenclamide group = 24/201 (12%) Insulin group = 26/203 (13%) p=0.76 Intravenous glucose therapy Glibenclamide group = 28/201 (14%) Insulin group = 22/203 (11%) p=0.36 Neonatal hypoglycaemia (<40mg/dl) Glibenclamide group = 18/201 (9 %) Insulin group = 12/203 (6%) 0.25 NICU Admission Glibenclamide group = 12/201 (6%) Insulin group = 14/203 (7%) p=0.68 Stillbirth Glibenclamide group = 1/201 (0.5%) Insulin group = 1/203 (0.5%) p=0.99 Neonatal death Glibenclamide group = 1/201 (0.5%) Insulin group = 1/203 (0.5%) p=0.99	manual. Appendix C: Methodology checklist: randomised controlled trials Appropriate randomisation method: Yes Adequate allocation concealment: Yes Groups comparable at baseline: Yes Groups received the same care (apart from the intervention): Yes Participants kept 'blind' to allocation: No Care givers kept 'blind' to allocation: No Follow up equal for groups: Yes How many participants did not complete treatment in each group?: None Were the groups were comparable for treatment completion: Yes For how many participants in each group were no outcome data available?: None The groups were comparable with respect to the

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td>funding
Not stated</td>
<td>y

Inclusion criteria
Women diagnosed with gestational diabetes (after screening using 50g GCT and a diagnostic 100g OGTT) who were attending maternal health clinics who had singleton pregnancies, were between 11-33 weeks gestation and who had FPG between 5.3mmol/l and 7.8 mmol/l at their diagnostic test.
Women with FPG <5.3mmol/l at their diagnostic test were initially treated with diet but were subsequently enrolled if their FPG ≥ 5.3mmol/l or the postprandial result was ≥ 6.7 mmol/l

Exclusion criteria
Not stated</td>
<td></td>
<td>tests to compare numerical data.</td>
<td></td>
<td>availability of outcome data: Yes
Appropriate length of follow-up: Yes
Precise outcome definitions used: Yes
Outcome determined using valid and reliable methods: Yes
Investigators kept 'blind' to allocation: Unclear
Investigators kept 'blind' to other important confounding and prognostic factors: Unclear

Other information
None.</td>
</tr>
<tr>
<td>Full citation
Louie,J.C., Markovic,T.P., Perera,N., Foote,D., Petocz,P., Ross,G.P., Brand-Miller,J.C., A randomized controlled trial investigating the effects of a low-glycemic index diet on pregnancy</td>
<td>Sample size
Total sample size comprised 99 women (7 were excluded leaving 92 women: 47 intervention, 45 control).

Characteristics
<table>
<tr><th>Characteristic</th><th>Low GI</th><th>Control</th><th>P-value</th></tr>
<tr><td>Mean age, years</td><td>34.0 ± 4.1</td><td>32.4 ± 4.5</td><td>0.06</td></tr>
<tr><td>Mean pre-</td><td>23.9</td><td>24.1 ±</td><td>0.84</td></tr>
</table>
</td>
<td>Interventions
Intervention
40 to 45% carbohydrate, 15 to 25% protein and 25 to 30% fat. A target GI of < 50 was imposed.

Control
40 to 45% carbohydrate, 15 to 25% protein and 25 to 30% fat.</td>
<td>Details
After diagnosis with GDM eligible women were randomised centrally using computer-generated random numbers strafified by BMI and gestational age.

At baseline and at 36 to 37 weeks' gestation women were asked to complete a three day food diary. This formed</td>
<td>Results
Large for gestational age
Low GI: 6/47
Control: 2/45
RR = 2.87 (95% CI 0.97 to 8.46)*

Emergency caesarean delivery
Low GI: 9/44
Control: 5/44
RR = 1.80 (0.64 to 1.85)*</td>
<td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to</td>
</tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
outcomes in gestational diabetes mellitus, Diabetes Care, 34, 2341-2346, 2011 Ref Id 177463 Country/ies where the study was carried out Australia Study type Randomised controlled trial. Aim of the study To determine the efficacy of a low glycaemic index diet versus a conventional healthy diet in reducing birth weight, birth weight centile, ponderal index and large for gestational age. Study dates September 2008 to November 2010.	<table><tr><td>pregnancy BMI, kg/m2</td><td>± 4.4</td><td>5.7</td><td></td></tr><tr><td>Ethnicity, %</td><td></td><td></td><td></td></tr><tr><td>Asian</td><td>59.6</td><td>55.6</td><td>0.70</td></tr><tr><td>Caucasian</td><td>31.9</td><td>40.0</td><td>0.42</td></tr><tr><td>Other</td><td>8.5</td><td>4.4</td><td>0.43</td></tr><tr><td>Mean fasting OGTT value, mmol/l</td><td>4.6 ± 0.5</td><td>4.7 ± 0.7</td><td>0.28</td></tr><tr><td>Mean 1 hour OGTT value, mmol/l</td><td>9.4 ± 1.4</td><td>9.7 ± 1.6</td><td>0.50</td></tr><tr><td>Mean 2 hour OGTT value, mmol/l</td><td>8.6 ± 1.2</td><td>8.0 ± 1.3</td><td>0.02</td></tr><tr><td>Nulliparous, %</td><td>61.7</td><td>64.4</td><td>0.79</td></tr></table> Inclusion criteria Aged 18 to 45, diagnosed with gestational diabetes mellitus by 75g OGTT at 20 to 32 weeks' gestation, healthy singleton pregnancy. Criteria for diagnosis of GDM were: Fasting glucose ≥ 5.5mmol/l 1 hour post-prandial glucose ≥ 10.0mmol/l 2 hour post-prandial glucose ≥ 8.0mmol/l Exclusion criteria Women with special dietary requirements (vegetarian/vegan), pre-existing diabetes,	A target GI of < 60 was imposed.	the basis of individualised dietary counselling. All participants received standard gestational diabetes care and were instructed in SMBG before breakfast and 1 hour after each meal. All participants and study staff were blinded to allocation, except the dietician. Large for gestational age was defined as birth weight > 90th percentile. Statistical analysis Based on a power of 80% to detect a 260g difference in birth weight. Primary analysis included women who attended at least one dietary session but excluded those with pre-term delivery (n = 4, 2 in each group). Pearson's X2 tests were used for categorical data. Continuous data were analysed using one-way ANOVA.	*Calculated by the NCC-WCH technical team.	treatment groups (which would have balanced any confounding factors equally across groups). Yes A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Yes A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Source of funding Funded by a grant from the Australian National Health and Medical Research Council.	pregnancy via assisted reproduction techniques and those who smoked or drank alcohol during pregnancy.		Paired t-tests were used to assess within-group changes from baseline. The study statistician was blinded to allocation.		B3: Individuals administering care were kept 'blind' to treatment allocation. Yes C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes - paired analysis for changes from baseline. C2: a. How many participants did not complete treatment in each group? 7 in total, groups not reported. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					treatment). Unclear C3: a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>outcome. Yes

D4: Investigators were kept 'blind' to participants' exposure to the intervention. Yes

D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Yes

Other information
None.</td></tr>
<tr><td>Full citation
Mesdaghinia,E., Samimi,M., Homaei,Z., Saberi,F., Moosavi,S.G., Yaribakht,M., Comparison of newborn outcomes in women with gestational diabetes mellitus treated with metformin or insulin: a randomised blinded trial, International</td><td>Sample size
N = 200

Characteristics
<table><tr><td>Characteristic</td><td>Insulin</td><td>Metformin</td></tr><tr><td>Mean maternal age, years</td><td>30.2 ± 5.9</td><td>29.6 ± 5.3</td></tr><tr><td>Mean BMI at start of pregnancy, kg/m2</td><td>28.46</td><td>27.60</td></tr><tr><td>Mean HbA_{1c}, %</td><td>6.3 ± 1.1</td><td>6.2 ± 1.6</td></tr><tr><td>Mean gestational age at randomisation, weeks</td><td>28.9 ± 3.8</td><td>27.9 ± 3.2</td></tr></table></td><td>Interventions
Intervention
Women in the metformin group received an initial dose of 500mg per day. If necessary this dose was adjusted up to a maximum of 2500g per day.

Control
Women in the insulin group received an initial dose of</td><td>Details
All women who met inclusion criteria were screened for GDM using a 1 hour 50g GCT. Women with impaired glucose tolerance based on the results of the GCT were given a 100g OGTT (one, two and three hours postprandial). Diagnosis of GDM was made if two abnormal values of the following were obtained:
Fasting glucose > 95mg/dl
1 hour postprandial > 180mg/dl</td><td>Results
Large for gestational age
Metformin: 16/100
Insulin: 24/100
RR = 0.67 (95% CI 0.05 to 8.51)*

Neonatal hypoglycaemia
Metformin: 10/100
Insulin: 15/100
RR = 0.67 (95% CI 0.32 to 1.42)*

NICU stay
Metformin: 14/100
Insulin: 33/100</td><td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Journal of Preventive Medicine, 4, 327-333, 2013 Ref Id 305965 Country/ies where the study was carried out Iran Study type Randomised controlled trial. Aim of the study To investigate outcomes in neonates of women treated with metformin compared with insulin. Study dates Not reported. Source of funding Not reported.	<table><tr><td>Family history of diabetes, n</td><td>12</td><td>9</td></tr></table> No comparisons were statistically significant. Specific p-values were not reported. Inclusion criteria Aged 18 to 45 years Singleton pregnancies No history of diabetes prior to pregnancy Gestational age 24 to 34 weeks Exclusion criteria Women treated with metformin who required supplemental insulin	0.5IU/kg/day (two thirds in the morning, one third in the afternoon). Two thirds of the insulin dose was NPH and one third regular insulin. One IU of insulin was added to the dose per 10mg/dl increase in blood glucose above target values.	2 hour postprandial > 155mg/dl 3 hour postprandial > 140mg/dl Women were randomised to receive either metformin (n = 100) or insulin (n = 100) using random number tables. Care providers and physicians assessing outcomes were blinded to allocation. Women were initially taught lifestyle modification and fasting and 2 hour postprandial blood glucose was measured for one week. If women obtained fasting values > 95mg/dl or 2 hour values > 120mg/dl pharmacological treatment was initiated. In the metformin group 22 out of 100 women randomised received supplemental insulin. These women were excluded and replaced. After achieving blood glucose targets women were discharged with a prescription and followed up every two weeks.	RR = 0.42 (95% CI 0.24 to 0.74)* Shoulder dystocia Metformin: 2/100 Insulin: 0/100 RR = 5.00 (95% CI 0.24 to 104.45)* *Calculated by the NCC-WCH technical team.	equally across groups). Yes A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Yes A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. No B3: Individuals administering care

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			Fasting and two hour postprandial blood glucose were recorded every two weeks until delivery and dosages adjusted accordingly. Outcomes included: LGA (not defined) NICU stay (definition not clear) Shoulder dystocia (not defined) Neonatal hypoglycaemia (not defined) Statistical analysis Sample size was calculated to have an 80% power to detect a difference of 0.13 between groups with a significance level of 0.05. It was not clear what the difference referred to but data were based on previous study results and the prevalence of GDM in Kashan city, Iran. Categorical data were analysed using either Fisher's exact test or the X2 test. Continuous data were analysed using either the Mann-Whitney U test or paired t-tests.		were kept 'blind' to treatment allocation. Yes C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? 22 women in the metformin group received insulin during the study therefore were excluded from analyses. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). No

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					C3: a. For how many participants in each group were no outcome data available? Not reported b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. No outcomes were defined. D3: A valid and reliable method was used to determine the

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					outcome. Unclear D4: Investigators were kept 'blind' to participants' exposure to the intervention. Yes D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information Women who failed treatment with metformin and required insulin were excluded from the study and replaced by women who had not failed treatment.
Full citation Moore,L.E., Briery,C.M., Clokey,D., Martin,R.W., Williford,N.J., Bofill,J.A., Morrison,J.C., Metformin and insulin in the management of gestational	Sample size 63 women were enrolled during 2001 to 2004 (Metformin group n=32, Insulin group n=31) Characteristics Characteristics / Metformin n=32 / Insulin n=32 / p value Age (year) / 27.1 ± 4.7 / 27.7 ± 6.7 / 0.778	Interventions All women received dietary instruction by a registered dietician and also from a nurse educator. The diet was designed to provide 30kcal/kg body weight. Women who were	Details Sample size calculations indicated that 128 participants (64 in each group) were required to achieve 80% power of detection of a significant ($p<0.05$) 10mg/dl difference in mean glucoe levels between the metformin and insulin groups. However, only 63	Results Metformin treatment failures (Definition: women who started taking insulin following 2 exceeded blood glucose targets over 2 consecutive weeks whilst receiving a maximum metformin dose 1000mg x 2/day) Metformin group = 0/32 27 women were controlled on the initial dose (500mg daily),	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials Appropriate randomisation method: Yes Adequate allocation concealment: Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
diabetes mellitus: preliminary results of a comparison, Journal of Reproductive Medicine, 52, 1011-1015, 2007 Ref Id 144586 Country/ies where the study was carried out United States of America Study type Randomised controlled trial Aim of the study To compare glycaemic control and neonatal outcomes in women diagnosed with gestational diabetes treated with metformin or insulin Study dates 2001 to 2004 at the University of Mississippi Medical Centre, Jackson Source of	<table><tr><td>Ethnicity (African American/ Native American/Caucasian)</td><td>20/11/1</td><td>11/17/3</td><td>0.087</td></tr><tr><td>Gravidity</td><td>3.1 ± 1.9</td><td>4.0 ± 2.5</td><td>0.171</td></tr><tr><td>Parity</td><td>1.4 ± 1.3</td><td>2.3 ± 2.3</td><td>0.173</td></tr><tr><td>Weight (kg)</td><td>104.28 ± 25.45</td><td>67.49 ± 19.5</td><td>0.01</td></tr><tr><td>Gestational age (weeks at study entry)</td><td>27.8 ± 6.5</td><td>28.9 ± 5.0</td><td>0.077</td></tr></table> Inclusion criteria Pregnant women received screening using 50g Glucose Challenge Test at 24-30 weeks gestation. Women who had levels >140mg/dl underwent a 3 hour diagnostic OGTT using ADA diagnostic criteria. Women with Class A2 gestational diabetes were defined as those who received dietary counselling and who failed to maintain fasting glucose <105mg/dl, and/or 2 hour postprandial glucose <120mg/dl. Women with Class 2 gestational diabetes were considered to require medication management. Women were eligible for inclusion if they had no renal or hepatic disease, hypertension or substance abuse histories. Exclusion criteria Not stated	obese (BMI>30) received a diet designed to deliver 25kcal/kg body weight. The calories were split by source: 40% carbohydrates, 20% protein, 30-40% fat. The patient received 10% at breakfast, 20-30% for both lunch and dinner and 30% for snacks.All women were trained to use a portable glucose meter at home and tested their blood glucose x3/day: in the morning (fasting value) and 2 hours after each meal. Metformin The initial dose was 500mg/day and was increased as necessary to attain glucose control (maximum dose 1000mg x2/day. Women taking the maximum dose of metformin with 2 values that	women had been recruited within the 32 month period and the results presented are an interim analysis of these participants' data. Randomisation and allocation to treatment group was performed using sequentially labelled, opaque sealed envelopes ordered by a computer generated list. After informed consent was obtained, a research nurse (not involved with patient care) selected the next envelope for the physician. Statistical analysis Sample size calculations were based on 80% power at the 0.05 significance level to detect a 10mg/dl difference in mean glucose levels between groups. The required sample size was 64 women per treatment arm. Student's t-test were used to compare means between groups. Fisher's exact tests were used to compare categorical data. Independent t-tests	4 women required a 1500mg/day dose and 1 woman required a 200mg/day dose Caesarean section Metformin group = 7/32 Insulin group = 10/31 p= 0.102 Birthweight > 4.0kg Metformin group = 3/32 Insulin group = 5/31 p=0.616 Birthweight > 4.5kg Metformin group = 0/32 Insulin group = 1/31 p= 0.321 NICU admission Metformin group = 2/32 Insulin group = 4/31 p=0.368 Neonatal hypoglycaemia (Definition: blood glucose <40mg/dl at 30 minutes or less after delivery) Metformin group = 0/32 Insulin group = 2/31 p=0.144 Shoulder dystocia Metformin group = 1/32 Insulin group = 0/31 p=0.321	Groups comparable at baseline: Yes except women in the metformin group were significantly heavier than those in the insulin group Groups received the same care (apart from the intervention): Yes Participants kept 'blind' to allocation: No, not possible Care givers kept 'blind' to allocation: No Follow up equal for groups: Yes How many participants did not complete treatment in each group?: None Were the groups were comparable for treatment completion: Yes For how many participants in each group were no outcome data available?: None The groups were comparable with respect to the availability of outcome data: Yes Appropriate length of follow-up: Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
funding Not stated		exceeded the goals for a measurement period for 2 consecutive weeks were considered metformin failures and were started on insulin. Insulin Insulin was started at a dosage of 0.7 units of insulin/kg actual body weight, and injected twice daily to maintain euglycaemia (fasting 60-90mg/dl; 2 hour postprandial <120mg/dl). The total daily dose was split; two thirds by subcutaneous injection in the morning and one third injected before the evening meal. A combination of regular insulin and NPH insulin was used.	and Mann-Whitney U tests were used where appropriate.		Precise outcome definitions used: Unclear for some outcomes Outcome determined using valid and reliable methods: Yes Investigators kept 'blind' to allocation: No Investigators kept 'blind' to other important confounding and prognostic factors: No Other information None.
Full citation Moore,L.E.,	Sample size N=149	Interventions Women were	Details Diet: All women were	Results Maternal outcomes	Limitations NICE guidelines

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Clokey,D., Rappaport,V.J., Curet,L.B., Metformin compared with glyburide in gestational diabetes: a randomized controlled trial, Obstetrics & Gynecology, 115, 55-59, 2010 Ref Id 145179 Country/ies where the study was carried out USA Study type Randomised controlled trial Aim of the study To compare the effects of metformin with glibenclamide on glycaemic control in women with gestational diabetes Study dates July 2003 and May 2008 Source of	An intention to treat analysis was performed Glibenclamide group = 74 (3 women did not take the treatment, 3 women relocated) Metformin group = 75 (5 women had only 2 prenatal visits, 2 women relocated, 1 woman only took 2 metformin doses due to gastrointestinal side effects) Characteristics Glibenclamide (n=74) / Metformin (n=75 / p value Hispanic: 66 / 66 / 0.81* Native American: 3 / 2 / White: 5 / 6 / African American: 0 / 1 / Age (yrs): 29.6± 7.8 / 31 ± 7.1 / 0.17 Weight (lbs): 180.1 ± 39 / 184.7 ± 35 / 0.49 Mean BMI: 32.7 ± 7.0 / 32.8 ± 5.8 / 0.88 BMI <30: 14 (19%) / 54 (72%) / BMI ≥ 30: 60 (81%) / 54 (72%) / Gestation at entry (wks): 29.1 ± 5.0 / 27.3 ± 6.8 / 0.10	randomised to treatment between 11 and 33 gestational weeks. Glibenclamide: An initial dose of 2.5mg twice per day was increased as necessary to a maximum dose of 20mg/day (10mg twice/day). Blood glucose was reviewed weekly. Metformin: An initial dose of 500mg/day taken in divided doses was increased as necessary to a maximum dose of 2grams/day. Blood glucose was reviewed weekly.	given instructions for a diet designed to provide 30kcals/kg at normal body weight and 25kcals/kg at obese body weight with 40% calories from carbohydrates, 20% from protein and 30-40% from fats.10% of calories were consumed at breakfast, 20-30% at lunch and dinner and 30% as snacks. Exercise: The importance of exercise in contolling blood glucose was stressed and 30 minutes of walking per day was recommended to all women. Monitoring: All women were taught how to use memory based glucometers. Women performed testing in the fasting state and 2 hours post prandially. Compliance was assessed by polling the meter at visits and by meetings with the diabetes educator at each visit when medication use, diet and exercise were reported by the women. Treatment failures were defined as women taking	Non-elective Caesarean delivery Glibenclamide group = 2/74 (1 failure to progress, 1 nonreassuring fetal status) Metformin group = 11/75 (3 breech presentations, 8 nonreassuring fetal status) p=0.02 Treatment failure Glibenclamide group = 12/74 Metformin group = 26/75 p=0.01 Maternal Hypoglycaemia (<60mg/dl) Glibenclamide group = 1/74 Metformin group = 2/75 p=0.56 Neonatal outcomes Neonatal hypoglycaemia (<40mg/dl) Glibenclamide group = 0/74 Metformin group = 1/75 p=0.32 Shoulder dystocia (no definition given) Glibenclamide group = 1/74 Metformin group = 0/75 p=0.49 NICU admission (no definition given) Glibenclamide group = 1/74 Metformin group = 4/75 p=0.37	manual. Appendix C: Methodology checklist: randomised controlled trials Appropriate randomisation method: Yes Adequate allocation concealment: Yes Groups comparable at baseline: Yes Groups received the same care (apart from the intervention): Yes Participants kept 'blind' to allocation: No Care givers kept 'blind' to allocation: No Follow up equal for groups: Yes How many participants did not complete treatment in each group?: Glibenclamide : 6 women Metformin : 8 women Were the groups comparable for treatment completion: Yes For how many participants in each group were no outcome data available?: None The groups were

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
funding Not stated	<table><tr><td>< 24 GW at entry</td><td>8 (11%)</td><td>13 (17%)</td><td>0.34</td></tr></table> * P value based on Hispanic compared with other Inclusion criteria Women were included if they had a diagnosis of gestational diabetes (using Carpenter and Coustan diagnostic criteria) and following diet and exercise counselling, did not maintain FPG < 105mg/dl or 2h postprandial blood glucose <120 mg.dl. Exclusion criteria A history of significant renal or hepatic disease, chronic hypertension requiring medication or substance misuse,		the maximum dose with two or more glucose values in the same meal exceeding target glucose values by 10mg/dl or more for 2 consecutive weeks. Oral medication was stopped in treatment failures and insulin therapy started. Statistical analysis The study was designed to have a power of 80% to detect a 10mg/dl difference in blood glucose between the two groups with a standard deviation of 20 mg/dl and an α = 0.5. Fisher's exact tests were used in the analysis of categorical data and Student's t-tests in the analysis of mean numerical data.		comparable with respect to the availability of outcome data: Yes Appropriate length of follow-up: Yes Precise outcome definitions used: No, Not for all outcomes Outcome determined using valid and reliable methods: Yes Investigators kept 'blind' to allocation: Unclear Investigators kept 'blind' to other important confounding and prognostic factors: Unclear Other information None.
Full citation Moreno-Castilla,C., Hernandez,M., Bergua,M., Alvarez,M.C., Arce,M.A., Rodriguez,K., Martinez-Alonso,M., Iglesias,M.,	Sample size N = 152 Characteristics <table><tr><td>Characteristic</td><td>Control</td><td>Low carbohydrate</td><td>P-value</td></tr><tr><td>Mean maternal</td><td>32.1 ± 4.4</td><td>33.5 ± 3.7</td><td>0.14</td></tr></table>	Interventions Intervention Low carbohydrate diet (40% of calories). Control Normal carbohydrate diet (55% of calories).	Details Women were screened for GDM between 24 and 28 weeks' gestation using a 50g GCT. If risk factors were present screening took place in the first trimester. A follow-up 100g OGTT was carried out on women with 1 hour GCT values ≥	Results Insulin treatment Low carbohydrate: 41/75 Control: 41/75 RR = 1.00 (95% CI 0.75 to 1.34)* Caesarean delivery Low carbohydrate: 25/74	Limitations NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual A. Selection bias A1: An appropriate

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Mateu,M., Santos,M.D., Pacheco,L.R., Blasco,Y., Martin,E., Balsells,N., Aranda,N., Mauricio,D., Low-carbohydrate diet for the treatment of gestational diabetes mellitus: a randomized controlled trial, Diabetes Care, 36, 2233-2238, 2013 Ref Id 309188 Country/ies where the study was carried out Spain Study type Randomised controlled trial. Aim of the study To assess whether a diet low in carbohydrates compared with a control diet could reduce the need for insulin treatment without increasing adverse	<table><tr><td>age, years</td><td></td><td></td><td></td></tr><tr><td>Mean pre-conception BMI, kg/m2</td><td>26.6 ± 5.5</td><td>25.4 ± 5.7</td><td>0.07</td></tr><tr><td>Mean gestational age at enrollment, weeks</td><td>30.1 ± 3.5</td><td>30.4 ± 3.0</td><td>0.89</td></tr><tr><td>Non-Caucasian, n (%)</td><td>6 (8.0)</td><td>1 (1.3)</td><td>0.12</td></tr><tr><td>Nulliparous, n (%)</td><td>37 (49.3)</td><td>40 (53.3)</td><td>0.74</td></tr></table> Inclusion criteria Aged 18 to 45 years Diagnosed with gestational diabetes mellitis Gestational age ≤ 35 weeks Exclusion criteria Unwillingness to follow a prescribed diet Inability to understand Spanish Pregnancy comorbidities other than obesity, hypertension or dyslipidaemia		7.8mmol/l. Diagnosis of GDM was made based on the Spanish National Diabetes Data Group criteria. A total of 152 women were randomised using sealed envelopes. Women were seen one week after allocation then subsequently every one to three weeks based on clinical judgement. All women were issued with a glucose meter and instructed to perform self-monitoring of blood glucose. All management strategies were the same for each group except for the intervention. Energy content of the diet was based on pregestational weight. Protein content of the diet was the same in each group (20%) but carbohydrate (40% intervention, 55% control) and fat (40% intervention, 25% control) differed. Diets were given as three meals and three snacks. No changes to the carbohydrate content of each diet were allowed	Control: 20/75 RR = 1.27 (95% CI 0.78 to 2.08)* Large for gestational age Low carbohydrate: 3/74 Control: 6/75 RR = 0.51 (95% CI 0.13 to 1.96)* Neonatal hypoglycaemia Low carbohydrate: 9/74 Control: 10/75 RR = 0.91 (95% CI 0.39 to 2.11)* *Calculated by the NCC-WCH technical team.	method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Unclear - used sealed envelopes but method of randomisation was not described. A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). No A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes B. Performance bias B1: The comparison groups received the same

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
outcomes. Study dates November 2008 to July 2011. Source of funding Not reported.			unless insulin therapy was initiated. Food records on 3 non-consecutive days including weekends and holidays were used to evaluate carbohydrate intake. Records were made after initial diet prescription and again after dietary plans were adjusted for adherence. Insulin therapy was initiated if at least two SMBG values in one week exceeded the following glycaemic targets: Fasting and preprandial ≤ 5.3mmol/l 1 hour postprandial ≤ 7.8mmol/l Neonatal hypoglycaemia was defined as < 2.2mmol/l. Large for gestational age was defined as birth weight > 90th percentile adjusted for sex and gestational age. Statistical analysis Sample size was calculated based on previous clinical data		care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. No B3: Individuals administering care were kept 'blind' to treatment allocation. No C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes C2: a. How many participants did not complete treatment in each group? One in each group (one before the intervention commenced, one after randomisation in the low carbohydrate group).

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			indicating that 40 to 50% of women with GDM require insulin treatment. The study was designed to provide 80% power to detect a 22% minimum difference for the risk of needing insulin therapy. The expected insulin therapy rate in the control group was 45%. Loss to follow-up was estimated to be 10%. A total sample size of 152 women (76 per arm) was calculated. Analyses were performed by a statistician blinded to allocation. Baseline characteristics were compared between groups to identify potential confounders. Results were analysed on an intention-to-treat basis with 95% confidence intervals and a significance level of 0.05.		b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Yes C3: a. For how many participants in each group were no outcome data available? One participant in the low carbohydrate group had no available data for Caesarean delivery, LGA and neonatal hypoglycaemia. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					whom outcome data were not available). Yes D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Yes D5: Investigators were kept 'blind' to other important confounding and prognostic factors. No Other information None.
Full citation	Sample size	Interventions	Details	Results	Limitations

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Moses,R.G., Barker,M., Winter,M., Petocz,P., Brand-Miller,J.C., Can a low-glycemic index diet reduce the need for insulin in gestational diabetes mellitus? A randomized trial, Diabetes Care, 32, 996-1000, 2009 Ref Id 145181 Country/ies where the study was carried out Australia Study type Randomised controlled trial. Aim of the study To determine whether a low glycaemic index diet in women with gestational diabetes reduces the need for insulin without compromising foetal or maternal	Total sample size comprised 63 women (31 intervention, 32 control). Characteristics Characteristic / Low GI / High GI / P-value Mean age, years / 30.8 ± 0.7 / 31.3 ± 0.8 / 0.68 Mean BMI at enrollment, kg/m2 / 32.0 ± 1.2 / 32.8 ± 1.4 / 0.68 Mean parity / 0.84 ± 0.17 / 0.78 ± 0.18 / 0.82 Mean fasting OGTT, mmol/l / 4.6 ± 0.1 / 4.7 ± 0.1 / 0.49 Mean 2 hour OGTT, mmol/l / 8.4 ± 0.2 / 8.4 ± 0.1 / 0.83 Inclusion criteria Aged 18 to 40 years, singleton pregnancy, no history of gestational diabetes, first clinical visit between 28 and 32 weeks' gestation and the ability to follow the study protocol requirements. Criteria for diagnosis of GDM using a 75g OGTT carried out at the start of the third trimester were: Fasting glucose ≥ 5.5mmol/l (100mg/dl) 2 hour post-prandial glucose ≥ 8.0mmol/l (145mg/dl)	Intervention Carbohydrate intake aimed to achieve a minimum of 175g per day. Foods included pasta, grain breads and unprocessed cereals with a high fibre content. Participants were told to avoid white bread, processed commercial cereals, potatoes and some types of rice. Control Carbohydrate intake aimed to achieve a minimum of 175g per day. Participants were advised to follow a high fibre and low-sugar diet. Whole wheat bread, potatoes and high fibre moderate-to-high GI breakfast cereals were recommended.	Eligible women potentially interested in participating were given a three day food diary between 28 and 32 weeks' gestation prior to assessment by a dietician. Visits 2 and 3 were 1 to 2 and 3 to 4 weeks after the first where 7 day food diaries were issued. Dieticians were not blinded. Women who agreed to participate were randomised using permuted blocks of unequal sizes generated using STATA. Insulin was initiated immediately to women in the low GI group if, more than once per week: Fasting glucose ≥ 5.5mmol/l, and/or 1 hour post-prandial glucose ≥ 8.0mmol/l Women in the high GI group were switched to the low GI diet if they exceeded these values. Large for gestational age was defined as > 90th percentile, adjusted for sex, gestational week of	Treatment failure Treatment: 9/31 Control: 19/32 RR = 0.49 (95% CI 0.26 to 0.91)* Large for gestational age Low GI: 3/31 High GI: 3/32 RR = 1.03 (95% CI 0.22 to 4.76)* *Calculated by the NCC-WCH technical team.	NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Yes A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear - attending physicians were not informed of allocation, dieticians were. No description of blinding of investigators. A3: The groups

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
outcomes. Study dates October 2007 to September 2008. Source of funding Funded by an internal revenue from the Illawarra Diabetes Service and the University of Sydney.	Exclusion criteria Any condition or medication which could affect glucose levels and refusal to follow the prescribed diet.		delivery, maternal age, parity, height and pre-pregnancy weight. Statistical analysis Independent t-tests were used to compare dietary components at different time points. Pearson X2 tests were used to compare proportions of participants requiring insulin with those who did not require insulin. P-values < 0.05 were considered to be significant.		were comparable at baseline, including all major confounding and prognostic factors. Yes B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. Unclear B3: Individuals administering care were kept 'blind' to treatment allocation. Yes C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2:

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					a. How many participants did not complete treatment in each group? None b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Yes C3: a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information 19 (59%) of the 32 women in the control arm required insulin

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>therefore were switched to a low GI diet during the trial.</td></tr>
<tr><td>Full citation
Mukhopadhyay,P., Bag,T.S., Kyal,A., Saha,D.P., Khalid,N., Oral hypoglycemic glibenclamide: Can it be a substitute to insulin in the management of gestational diabetes mellitus? a comparative study, Journal of SAFOG, 4, 28-31, 2012
Ref Id
236621
Country/ies where the study was carried out
India
Study type
Randomised controlled trial.

Aim of the study
To compare insulin with glibenclamide for</td><td>Sample size
N = 60

Characteristics
<table><tr><th>Characteristic</th><th>Glibenclamide</th><th>Insulin</th></tr><tr><td>Mean maternal age, years</td><td>26.3 ± 4.6</td><td>26.0 ± 4.3</td></tr><tr><td>Mean BMI, kg/m2</td><td>23.7 ± 2.7</td><td>23.0 ± 2.9</td></tr><tr><td>Mean gestational age at entry, weeks</td><td>28.3 ± 2.2</td><td>27.4 ± 2.7</td></tr></table>
P-values were not reported.

Inclusion criteria
Diagnosis of GDM
20 to 28 weeks' gestation
Singleton pregnancies

Exclusion criteria
Women with pre-existing diabetes
Severe anaemia
Heart diseases
Renal disorders
Women taking steroids</td><td>Interventions
Intervention
The initial dose of glibenclamide was 2.5mg orally in the morning. Doses were increased when necessary by 2.5mg per week up to a maximum of 20mg per week. Doses > 7.5mg were given as divided doses. If glycaemic control was not maintained for two weeks on the maximal dose then treatment was switched to insulin.

Control
Insulin treatment was initiated at 0.7units/kg/day, subcutaneously three times daily and increased weekly as necessary.</td><td>Details
Women attending the antenatal clinic of the study hospital were screened for GDM using a 75g oral glucose. Diagnosis of GDM was made based on 2 hour postprandial values > 140mg/dl according to the WHO criteria.

Women who met inclusion criteria were given nutritional therapy for two weeks. Caloric intake was calculated according to BMI. A total of 60 women did not achieve glycaemic control using dietary therapy. The goal of treatment was fasting glucose < 90mg/dl and postprandial peaks < 120mg/dl. The 60 women were randomised to either glibenclamide (n = 30) or insulin (n = 30) using random number tables.

Women were instructed to self-monitor blood</td><td>Results
Large for gestational age
Glibenclamide: 4/30
Insulin: 2/30
RR = 2.00 (95% CI 0.38 to 10.45)*

Neonatal hypoglycaemia
Glibenclamide: 4/30
Insulin: 3/30
RR = 1.33 (95% CI 0.32 to 5.60)*

*Calculated by the NCC-WCH technical team using the t-distribution due to small sample size.</td><td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Yes

A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear

A3: The groups</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
the treatment of gestational diabetes mellitus. Study dates January 1st to December 31st 2010. Source of funding Not reported.			glucose seven times daily. Laboratory measurements were also taken each week. Outcomes included: Large for gestational age (birth weight > 90th percentile) Neonatal hypoglycaemia (< 44mg/dl) Statistical analysis Data between groups were compared using the Student's t-test.		were comparable at baseline, including all major confounding and prognostic factors. Unclear - minimal baseline characteristics were reported. B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. No B3: Individuals administering care were kept 'blind' to treatment allocation. Unclear C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up).

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					Unclear C2: a. How many participants did not complete treatment in each group? Not reported b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Unclear C3: a. For how many participants in each group were no outcome data available? Not reported b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>None</td></tr>
<tr>
<td>Full citation
Niromanesh,S., Alavi,A., Sharbaf,F.R., Amjadi,N., Moosavi,S., Akbari,S., Metformin compared with insulin in the management of gestational diabetes mellitus: A randomized clinical trial, Diabetes Research and Clinical Practice, 98, 422-429, 2012
Ref Id
248270
Country/ies where the study was carried out
Iran
Study type
Randomised controlled trial.

Aim of the study
To evaluate the effect of metformin and insulin in glycaemic control in pregnant</td>
<td>Sample size
N = 172

Characteristics
<table>
<tr><th>Characteristic</th><th>Metformin</th><th>Insulin</th><th>P-value</th></tr>
<tr><td>Mean maternal age, years</td><td>30.7 ± 5.5</td><td>31.8 ± 5.1</td><td>0.22</td></tr>
<tr><td>Mean BMI, kg/m2</td><td>28.1 ± 4.0</td><td>27.1 ± 2.1</td><td>0.06</td></tr>
<tr><td>Mean gestational age at entry, weeks</td><td>28.7 ± 3.7</td><td>28.6 ± 3.6</td><td>0.86</td></tr>
<tr><td>Mean HbA_{1c} at entry, %</td><td>5.7 ± 0.6</td><td>5.6 ± 0.7</td><td>0.59</td></tr>
<tr><td>Multipara, n (%)</td><td>12 (15.0)</td><td>16 (20.0)</td><td>0.69</td></tr>
<tr><td>History of macrosomia, n (%)</td><td>2 (2.5)</td><td>5 (6.3)</td><td>0.44</td></tr>
</table>

Inclusion criteria
Aged 18 to 40 years
Singleton pregnancies
Gestational age between 20 and 34 weeks
Blood glucose values > 95mg/dl fasting and > 120mg/dl 2 hour postprandial after nutritional therapy</td>
<td>Interventions
Intervention
Metformin was given as an initial dose of 500mg twice daily and increased by 500 to 1000mg up to a maximum dose of 2500mg divided dose with each meal. Metformin was continued until delivery. Insulin was added if glucose control was not achieved with maximal metformin doses.

Control
Women in the insulin group were treated with NPH insulin at an initial dose of 0.2units/kg. If fasting glucose was high insulin was given before bedtime. If postprandial glucose was high, regular short-acting insulin was given before meals based on postprandial</td>
<td>Details
All pregnant women receiving prenatal care at the study hospital were screened using a 50g GCT. Women with 1 hour glucose ≥ 130mg/dl were given a 3 hours 100g OGTT. Women with two or more abnormal values using Coustan and Carpenter's criteria were diagnosed with GDM.

All women were given counselling on diet and physical activity. Daily caloric intake was based on BMI. Carbohydrate intake was restricted to 45% of calories with remainder as protein (20%) and fat (35%). An exercise program of 30 minutes per day was recommended.

A total of 172 women meeting inclusion criteria were enrolled. Women with GDM inadequately controlled by diet were allocated to either metformin (n = 86) or insulin (n = 86) using sequentially labelled sealed envelopes numbered by a computer</td>
<td>Results
Shoulder dystocia
Metformin: 2/80
Insulin: 4/80
RR = 0.5 (95% CI 0.1 to 2.6)

Caesarean section
Metformin: 34/80
Insulin: 37/80
RR = 0.7 (95% CI 0.2 to 2.2)

Emergency Caesarean section
Metformin: 25/80
Insulin: 16/80
RR = 1.6 (95% CI 0.9 to 2.7)

Large for gestational age
Metformin: 14/80
Insulin: 28/80
RR = 0.5 (95% CI 0.3 to 0.9)

NICU stay
Metformin: 5/80
Insulin: 2/80
RR = 2.5 (95% CI 0.5 to 12.5)

Neonatal hypoglycaemia
Metformin: 3/80
Insulin: 2/80
RR = 1.5 (95% CI 0.3 to 8.7)

Treatment failure</td>
<td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Yes

A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Yes

A3: The groups were comparable at baseline, including all major confounding and</td>
</tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
women with GDM in relation to pregnancy outcomes. Study dates December 2010 to January 2012. Source of funding Not reported.	Exclusion criteria History of systemic underlying diseases (cardiovascular, renal, liver or autoimmune) Substance abuse Overt diabetes mellitus (except previous history of GDM) Major fetal malformations	glucose levels (1 unit for every 10mg/dl glucose). If both fasting and postprandial values were high insulin was started at a dose of 0.7units/kg (two thirds NPH insulin before breakfast and bedtime, one third regular insulin as two or three preprandial injections).	generated random number list. Obstetricians responsible for clinical and prenatal care were blinded to allocation. Women were instructed in the use of capillary glucose monitoring by a nurse. SMBG was to be undertaken four times per day. Target blood glucose values were as follows: Fasting glucose < 95mg/dl Postprandial (no time given) < 120mg/dl Women were asked to participate if 2 readings were abnormal based on self-assessment. Women then monitored blood glucose bi-weekly. Pharmacological treatment was started if two fasting, one fasting and one postprandial or two postprandial values were above the glucose targets. Primary study outcomes were maternal glycaemic control and birth weight.	Metformin: 11/80 Insulin: not reported RR not calculable	prognostic factors. Yes B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. No B3: Individuals administering care were kept 'blind' to treatment allocation. Yes C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes C2: a. How many participants did not complete treatment in each group? Out

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			Elective delivery was planned for 38.5 weeks' gestation by induction of labour or Caesarean. Other outcomes included: Shoulder dystocia (not defined.) Admission to NICU (not defined) Macrosomia (birth weight ≥ 4000g) LGA (birth weight > 90th percentile) Perinatal death (not defined) Neonatal hypoglycaemia (not defined) Mode of birth (overall and emergency Caesarean) Statistical analysis Sample size was calculated to provide a power of 85% to detect a 225g difference in birth weight between groups with a standard deviation of 450g and to detect a 10mg/dl difference in blood glucose with a standard deviation of 20mg/dl. The significance level was set at 0.05. Continuous variables were compared between		of 86 women in each group: two were lost to follow-up and three discontinued treatment due to side effects in the metformin group; six were lost to follow-up but none discontinued treatment in the insulin group. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). No C3: a. For how many participants in each group were no outcome data available? b. The groups were comparable with respect to the availability of outcome data (that is, there were no

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			groups using independent sample t-tests. Categorical variables were compared using the χ2 test or Fisher's exact test. Relative risks and 95% confidence intervals were calculated. Binary logistic regression was performed to determine predictors of LGA.		important or systematic differences between groups in terms of those for whom outcome data were not available). Yes / no / unclear / N/A D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. No - shoulder dystocia, NICU stay, perinatal death and neonatal hypoglycaemia were not defined. D3: A valid and reliable method was used to determine the outcome. Unclear D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information None
Full citation Ogunyemi,D., Jesse,M., Davidson,M., Comparison of glyburide versus insulin in management of gestational diabetes mellitus, Endocrine Practice, 13, 427-428, 2007 Ref Id 155679 Country/ies where the study was carried out United States of America Study type Open label randomised controlled trial Aim of the study To compare the	Sample size 97 women randomised to treatment with glibenclamide (n=48) or insulin (n=49) Characteristics 80% of participants were Hispanic and 15% were African American. The treatment groups were similar at baseline for maternal age, parity, BMI, history of previous gestational diabetes and precious neonatal macrosomia. Results of the 1 hour 50g GCT, HbA_{1c} testing and 3 hour OGTT (fasting, 1 hour and 2 hour post load values) were significnatly higher in the insulin group compared to the glibenclamide group. The gestational age at the time of recruitment to the study was 4 weeks later in the glibenclamide group compared to the insulin group. Inclusion criteria Diet therapy had not been successful in all participants. No other details are presented Exclusion criteria No details are presented	Interventions No diet or monitoring details are presented No details of dose for glibenclamide or insulin are presented	Details Randomisation was performed using a computer generated list and treatment assignation was performed using sequentially numbered opaque sealed envelopes. Statistical analysis Not reported.	Results Treatment failure Glibenclamide = 3/48 women were transitioned to insulin Maternal hypoglycaemia Glibenclamide = 18/48 (38%) Insulin = 15/49 (31%) Caesarean delivery Glibenclamide = 18/43 (42%) Insulin = 25/45 (56%) Neonatal hypoglycaemia Glibenclamide = 12/43 (28%) Insulin = 6/45 (13%) Birth defects Glibenclamide = 4/43 (9%) Insulin = 3/45 (7%)	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials Appropriate randomisation method: yes Adequate allocation concealment: yes Groups comparable at baseline: no Groups received the same care (apart from the intervention): unclear, not stated Participants kept 'blind' to allocation: no Care givers kept 'blind' to allocation: no Follow up equal for groups: yes How many

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
effects of glibenclamide with insulin on maternal glucose control and neonatal outomes in women with gestational diabetes. Study dates 2002 to 2005 Source of funding Not stated					participants did not complete treatment in each group?: none Were the groups comparable for treatment completion: yes For how many participants in each group were no outcome data available?: Up to 4 in the insulin group and 5 in the glibenclamide group The groups were comparable with respect to the availability of outcome data: unclear, not stated Appropriate length of follow-up: yes Precise outcome definitions used: unclear, not stated Outcome determined using valid and reliable methods: yes Investigators kept 'blind' to allocation: no Investigators kept 'blind' to other important confounding and prognostic factors:no

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Other information
None.</td></tr>
<tr><td>Full citation
Persson,B., Stangenberg,M., Hansson,U., Nordlander,E., Gestational diabetes mellitus (GDM). Comparative evaluation of two treatment regimens, diet versus insulin and diet, Diabetes, 34 Suppl 2, 101-105, 1985
Ref Id
177572
Country/ies where the study was carried out
Sweden
Study type
Randomised controlled trial.

Aim of the study
To compare the effect of diet plus insulin with diet alone on</td><td>Sample size
Total sample size comprised 202 women (97 intervention, 105 control).

Characteristics<table><tr><td>Characteristic</td><td>Diet alone</td><td>Diet + insulin</td></tr><tr><td>Median age, years (IQR)</td><td>29 (18 to 46)</td><td>30.5 (16 to 42)</td></tr><tr><td>Median pre-pregnancy weight, kg (IQR)</td><td>60 (44 to 130)</td><td>64.7 (39 to 120)</td></tr><tr><td>Parity = 0, n</td><td>32</td><td>27</td></tr><tr><td>Parity ≥ 1, n</td><td>73</td><td>70</td></tr></table>
No significant differences were observed. P-values were not reported.

Inclusion criteria
OGTT area under the curve of ≥ 2 SD above normal after a 3 hour 50g OGTT.

Exclusion criteria
Not reported.</td><td>Interventions

Intervention

Diet plus an initial dose of 8 to 12IU/day of intermediate or fast-acting insulin.

Control
Diet comprising 50% calories from carbohydrates, 20% from protein, 30% from fat.</td><td>Details
239 women met inclusion criteria. Of these 37 women refused to participate leaving 202 who were randomised to either diet plus insulin or diet alone.

All women were given dietary advice by a dietician and instructed to follow the prescribed diet. All participants were instructed in SMBG which was carried out on 3 days per week, 6 times each day.

If fasting or 1 hour post-prandial glucose exceeded 7mmol/l or 9mmol/l, respectively, ≥ 3 times in one week diet was deemed insufficient and insulin therapy initiated.

Outcomes included:
Large for gestational age (> 90th percentile for gestational age)
C-peptide concentration</td><td>Results

Treatment failure

Treatment: 15/105

Control: not reported

RR = not calculable

Large for gestational age
Diet + insulin: 11/97
Diet alone: 14/105
RR = 0.85 (95% CI 0.41 to 1.78)*

Hypoglycaemia
Diet + insulin: 20/97
Diet alone: 13/105
RR = 1.67 (95% CI 0.88 to 3.17)*

Perinatal mortality
Diet + insulin: 0/97
Diet alone: 0/105
RR not calculable.

C-peptide concentration (hyperinsulinaemia)
Data were presented as a</td><td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Unclear - stratified selection but sequence generation is not described.

A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
maternal glucose and neonatal outcomes in the treatment of women with gestational diabetes mellitus. Study dates November 1981 to May 1984. Source of funding Supported by grants from the Swedish Medical Research Council, the Tielman Fund for Pediatric Research, the Expression Fund for Prenatal Research, Almanna Minnesfond and the Swedish Diabetic Association.			Hypoglycaemia (not defined) Statistical analysis Between-group comparisons were made using ANOVA, X2 tests or Mann-Whitney U tests. Women who "failed" diet alone treatment (required insulin) were included in analyses.	figure therefore analysis was not possible. *Calculated by the NCC-WCH technical team.	allocation). Unclear A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. N/A B3: Individuals administering care were kept 'blind' to treatment allocation. N/A C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up).

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					Unclear C2: a. How many participants did not complete treatment in each group? 1 in the diet + insulin group, none in the diet alone group. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Yes C3: a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. No - hypoglycaemia not defined. D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Other information
None.</td></tr>
<tr><td>Full citation
Rae,A., Bond,D., Evans,S., North,F., Roberman,B., Walters,B., A randomised controlled trial of dietary energy restriction in the management of obese women with gestational diabetes, Australian and New Zealand Journal of Obstetrics and Gynaecology, 40, 416-422, 2000
Ref Id
177595
Country/ies where the study was carried out
Australia
Study type
Randomised controlled trial.</td><td>Sample size
Total sample size comprised 125 women, 8 withdrew (63 intervention, 54 control).

Characteristics
<table><tr><th>Characteristic</th><th>Intervention</th><th>Control</th><th>P-value</th></tr><tr><td>Mean age</td><td>30.2</td><td>30.6</td><td>0.66</td></tr><tr><td>Nulliparity, n</td><td>18</td><td>17</td><td>0.73</td></tr><tr><td>Mean BMI at diagnosis</td><td>37.9 ± 0.7</td><td>38.0 ± 0.7</td><td>0.90</td></tr></table>
Inclusion criteria
Gestation ≤ 35 weeks and 6 days, > 110% ideal body weight and a positive OGTT test result.

Criteria for diagnosis by OGTT were:
Fasting glucose > 5.4mmol/l and/or
2 hour plasma glucose > 7.9mmol/l

Exclusion criteria
Not reported.</td><td>Interventions
Intervention
Instruction in a moderately energy-restricted diet comprising 1590 to 1776kcal per day (70% of the RDI for pregnant women).

Control
Instruction in an unrestriced diabetic diet comprising 2010 to 2220kcal per day.</td><td>Details
Eligible women were randomised according to strata of maternal age, gestational age at diagnosis, parity and the degree of abnormality of the OGTT results. Randomisation was carried out by drawing sealed numbered envelopes. Participants and clinical staff were blinded to allocation.

Moderate GDM was defined as fasting plasma glucose between 5.5 to 5.8mmol/l or 2 hour post-prandial blood glucose between 8.0 to 8.9mmol/l. Severe GDM was defined as any one measurement above these values or both fasting and 2 hour values above the thresholds for GDM (see interventions section).

All participants received education, control of hyperglycaemia and</td><td>Results
Induction of labour, n/N
Energy-restricted diet: 29/63
Control: 23/51
RR = 1.02 (95% CI 0.18 to 5.76)*

Vaginal delivery (spontaneous), n/N
Energy-restricted diet: 31/65
Control: 30/56
RR = 0.89 (95% CI 0.63 to 1.27)*

Caesarean delivery, n/N
Energy-restricted diet: 26/65
Control: 19/56
RR = 1.18 (95% CI 0.74 to 1.89)*

Treatment failure
Treatment: 11/63
Control: 9/54
RR = 1.05 (95% CI 0.47 to 2.34)*

Shoulder dystocia
Energy-restricted diet: 0/65
Control: 0/56</td><td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Unclear - method of numbering envelopes is not described.

A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Aim of the study To determine whether moderate energy restriction would reduce the need for insulin in women with gestational diabetes mellitus and the incidence of macrosomia. Study dates February 1992 and June 1995. Source of funding Supported by a grant from the Foundation for Women's and Infant's Health, Western Australia.			foetal and maternal surveillance. Insulin therapy was started if, on ≥ 2 occasions: Fasting glucose > 5.5mmol/l 2 hour post-prandial glucose > 7.0mmol The decision to start insulin was made by clinical staff blinded to allocation. SMBG was performed before and 2 hours after each meal at least two days per week. Compliance was monitored using three day food diaries at three time periods. "Macrosomia" (LGA) was defined as > 90th percentile for gender, gestational age and maternal height. Statistical analysis Baseline data were compared using Student's t-tests, Wilcoxon rank-sum tests or Fisher's exact tests. Factors affecting insulin use and macrosomia	RR not calculable. *Calculated by the NCC-WCH technical team.	or treatment allocation). Yes A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. Yes B3: Individuals administering care were kept 'blind' to treatment allocation. Yes C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			rates were assessed using logistic regression. All other outcomes were analysed using multivariate repeated measures or linear ANOVA. Sample size was calculated to have 80% power to detect a reduction in insulin use from 40 to 15% and a reduction in macrosomia rates from 25 to 5%. Type 1 error was 0.05. This provided a required sample size of 60 patients per group. Data were analysed on an intention to treat basis.		length of follow-up). Yes - repeated measures analysis was used. C2: a. How many participants did not complete treatment in each group? 4 in each group. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Unclear C3: a. For how many participants in each group were no outcome data available? Not reported - denominators not reported for several outcomes, frequencies only. b. The groups were comparable with

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. No - shoulder dystocia was not defined. D3: A valid and reliable method was used to determine the outcome. Unclear D4: Investigators were kept 'blind' to participants' exposure to the intervention. Yes D5: Investigators were kept 'blind' to

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					other important confounding and prognostic factors. Unclear Other information None.
Full citation Rowan,J.A., Hague,W.M., Gao,W., Battin,M.R., Moore,M.P., MiG,Trial,I, Metformin versus insulin for the treatment of gestational diabetes.[Erratum appears in N Engl J Med. 2008 Jul 3;359(1):106], New England Journal of Medicine, 358, 2003-2015, 2008 Ref Id 145223 Country/ies where the study was carried out Australia Study type Open-label,	Sample size The study was conducted in 10 New Zealand and Australian urban obstetric hospitals. Of the 751 women recruited to the study, the analyses included 363 women in the metformin group and 370 in the insulin group (n=733) and were performed according to the intention-to-treat principle. Data after randomization were not available for 10 women in the metformin group and 8 in the insulin group. Characteristics The two groups were similar at baseline for 21 characteristics including age, BMI, gestation length at enrollment, race/ethnic group, smoking,blood pressure, diagnostic test result and obsteric and family history parameters. However, more women in the metformin group than in the insulin group had had 3 or more pregnancy terminations or miscarriages (23.1% vs 16.8%, p = 0.03). Inclusion criteria Women were eligible for inclusion if they were between 18 and 45 years of age, had received a diagnosis of gestational diabetes mellitus according to ADIPS 1998 criteria,	Interventions All women received some lifestyle advice about diet and exercise prior to randomisation. All sites aimed for ADIPS 1998 recommendations for capillary glucose levels (fasting <5.5 mmol/l; 2-hour postprandial <7.0 mmol/l), several sites aimed for lower target levels. Metformin Local pharmacies supplied medications to women according to prescription. Metformin was supllied as Metomin [Pacific Pharmaceuticals]	Details The primary aim of the study was to rule out a clinically significant increase (from 30% to 40%) of the primary composite outcome in the metformin group. The anticipated rates for each component were 14% for hypoglycemia, 5% for respiratory distress, 5% for phototherapy, 1.5% for birth trauma, < 1% for Apgar scores below 7, and 15% for preterm delivery. The infants could meet one or more of the criteria. Two-tailed calculations were used to rule out a significant difference in either direction. For 80% power and a 5% significance level, 375 subjects were required in each group. Block randomisation was performed with stratification according to	Results Maternal outcomes Induction of labor Metformin group = 196 women (54.0%) Insulin group = 208 (56.2%) (P = 0.55) Cesarean section Metformin group = 131 women (36.1%) Insulin group = 142 (38.4%) (P = 0.52) Emergency cesarean section Metformin group = 55 women (15.2%) Insulin group = 63 (17.0%) (P = 0.49) Treatment failure Supplemental insulin was required in 168 women (46.3%) in the metformin group. Metformin treatment was stopped in 27 women (7.4%) before delivery (Fig. 1). Treatment was stopped in 11 of these women in accordance with the trial protocol (9 women had obstetrical	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials Appropriate randomisation method: Yes Adequate allocation concealment: Unclear Groups comparable at baseline: Yes except for one point of obstetrc history Groups received the same care (apart from the intervention): Yes Participants kept 'blind' to allocation: No Care givers kept 'blind' to allocation: No Follow up equal for groups: Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
randomised controlled trial Aim of the study The Metformin in Gestational Diabetes(MiG) Trial was designed to rule out a 33% increase in a composite of perinatal complications in infants of women treated with metformin as compared with insulin. The hypotheses were that perinatal outcomes would be similar for both treatments, that women would consider metformin a more acceptable treatment than insulin, and that metformin would improve markers of insulin sensitivity in the mother and baby. Study dates October 2002 and November	were pregnant with a single fetus between 20 and 33 weeks of gestation, met the hospital's usual criteria for starting insulin treatment, and, after lifestyle intervention consisting of advice about diet and exercise, had more than one capillary blood glucose measurement above 5.4 mmol/l after an overnight fast or more than one 2-hour postprandial blood glucose measurement above 6.7 mmol/l. Exclusion criteria Exclusion criteria were a prepregnancy diagnosis of diabetes, a contraindication to metformin, a fetal anomaly, gestational hypertension, preeclampsia, fetal growth restriction and ruptured membranes.	in New Zealand and as Diaformin Alphapharm and other nonspecified manufacturers] in Australia). The initial dose was 500 mg once or twice daily with food and was typically increased over 1 to 2 weeks, to meet glycemic targets up to a maximum daily dose of 2500 mg. If the targets were not achieved with metformin alone, insulin was added. Metformin was stopped if maternal contraindications (such as liver or renal impairment or sepsis) or fetal growth restriction developed. Insulin Insulin was prescribed according to usual practice.	site and gestational age (from 20 to 27+6 weeks or from 28 to 33+6 weeks). The primary outcome was a composite of the following neonatal complications: neonatal hypoglycemia (two or more neonatal glucose values <2.6 mmol per litre), respiratory distress (need for at least 4 hours of respiratory support with supplemental oxygen, continuous positive airway pressure, or intermittent positive-pressure ventilation during the first 24 hours after delivery), need for phototherapy, birth trauma (injury to the baby at delivery, documented as mild if bruises or abrasions were present at birth but resolved before 6 weeks post partum; more serious injuries were also recorded), 5-minute Apgar score below 7, or premature birth (<37 weeks of gestation). The component complications were chosen to reflect important adverse effects of fetal exposure to maternal hyperglycemia that might be modified by	complications, 1 had sepsis, and 1 had worsening abnormal liverfunction test results); treatment was stopped in 7 women (1.9%) because of gastrointestinal side effects; 5 women chose to stop metformin; and 4 women were advised to stop by other health professionals who were not involved in the trial. Metformin doses were reduced because of gastrointestinal side effects in 32 women (8.8%); all but 1 of these women were able to maintain a dose of at least 1000 mg per day. Results of questionnaire regarding acceptability of treatment How often did you forget to take your medication? p < 0.001 Never or rarely: Metformin Group = 231/333 (69.4%) Insulin Group = 267/331 (80.7%) 1–3 times/wk: Metformin Group = 81/333 (24.3%) Insulin Group = 52/331 (15.7%) 4–6 times/wk: Metformin Group = 12/333 (3.6%) Insulin Group = 2/331 (0.6%) >6 times/wk: Metformin Group = 9/333 (2.7%) Insulin Group = 10/331 (3.0%)	How many participants did not complete treatment in each group?: Metformin group =27, Insulin Group = 0 Were the groups were comparable for treatment completion: Only repotred for metformin group For how many participants in each group were no outcome data available?: Data after randomization were not available for 10 women in the metformin group and 8 in the insulin group. The groups were comparable with respect to the availability of outcome data: Yes Appropriate length of follow-up: Yes Precise outcome definitions used: Yes Outcome determined using valid and reliable methods: Yes Investigators kept 'blind' to allocation: No

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
2006 Source of funding Grants from the Auckland Medical Research Foundation, the National Women's Evelyn Bond Charitable Trust, the Health Research Council of New Zealand, and the National Health and Medical Research Council of Australia			treatment and directly influenced by metformin crossing the placenta. Neonates were monitored for hypoglycemia by measuring blood glucose levels within 2 hours after birth and before each feeding until consecutive glucose values of 2.6 mmol per liter or greater were achieved. Readings below 2.6 mmol per liter and below 1.6 mmol per liter were documented, as was treatment for hypoglycemia. A questionnaire was administered to the mothers in the first postpartum week to assess acceptability of the treatment as a secondary outcome measure. Adverse events were reported to the data and safety monitoring committee. Side effects of medication and complications of pregnancy were documented at clinic visits, and the investigators were informed of hospitalizations. Congenital anomalies and events that were fatal, life-threatening,	Which medication would you choose in another pregnancy? p < 0.001 Metformin tablets: Metformin Group = 256/334 (76.6%) Insulin Group = 127/331 (38.4%) Insulin injections: Metformin Group = 42/334 (12.6%) Insulin Group = 90/331 (27.2%) Not sure: Metformin Group = 36/334 (10.8%) Insulin Group = 114/331 (34.4%) In another pregnancy, if you were told you were likely to need insulin injections to control the sugar levels but could try metformin first, what would you prefer? p < 0.001 Start with metformin and add insulin if needed: Metformin Group = 270/334 (80.8%) Insulin Group = 179/331 (54.1%) Go straight to insulin injections: Metformin Group = 36/334 (10.8%) Insulin Group = 94/331 (28.4%) Not sure: Metformin Group = 28/334 (8.4%) Insulin Group = 58/331 (17.5%) Which part of your diabetes treatment was the easiest? p < 0.001 Doing finger-prick tests: Metformin Group = 74/334 (22.2%) Insulin Group = 119/331 (36.0%) Being careful with diet: Metformin Group = 63/334	Investigators kept 'blind' to other important confounding and prognostic factors:No Other information None.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			associated with serious disability or incapacity, required prolonged hospitalization (apart from hospitalization related to expected pregnancy events), or required a major intervention to prevent another serious outcome were classified as serious adverse events. Other measures of neonatal complications were admission to a level 2 or level 3 neonatal intensive care unit, duration of stay in the neonatal intensive care unit, and diagnosis at discharge from the hospital. Statistical analysis The study was powered to rule out a clinically significant increase in the primary outcome of 30 to 40% in the metformin group. Based on 80% power and a significance level of 0.05 the required sample size was 375 women in each arm. Between-group differences were analysed using X2 tests or Fisher's exact tests where appropriate. Two-sample t-tests or Mann-	(18.9%) Insulin Group = 95/331 (28.7%) Taking medication: Metformin Group = 197/334 (59.0%) Insulin Group = 117/331 (35.3%) Which part of your diabetes treatment was the hardest? p = 0.001 Doing finger-prick tests: Metformin Group = 123/334 (36.8%) Insulin Group = 91/331 (27.5%) Being careful with diet: Metformin Group = 176/334 (52.7%) Insulin Group = 150/331 (45.3%) Taking medication: Metformin Group = 35/334 (10.5%) Insulin Group = 90/331 (27.2%) Neonatal outcomes Birth weight >90th percentile Birth-weight percentiles were calculated with the use of a customized calculator that adjusts for sex and gestational age of the infant, as well as maternal height, weight in early pregnancy, ethnic group, and parity Metformin Group = 70/363 (19.3%) Insulin Group = 69/370 (18.6%) p=0.83 >24hour stay in NICU Metformin Group = 46/363 (12.7%) Insulin Group = 45/370 (12.2%)	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			Whitney U tests were used to analyse continuous data. Interim analyses were carried out. P-values were adjusted using the Peto-Haybrittle method. Investigators were to be informed when a between-group difference ≥ 3 standard deviations was observed.	Relative Risk (95% CI) = 1.04 (0.71–1.53) p = 0.83 Primary composite outcome Metformin Group = 116/363 (32.0%) Insulin Group = 119/370 (32.2) Relative Risk (95% CI) = 0.99 (0.80–1.23) p = 0.95 Supplemental feeding Metformin group =129 infants (35.5%) Insulin group = 145 (39.2%) p = 0.31 Intravenous dextrose Metformin group = 25 infants (6.9%) Insulin group = 22 (5.9%) p = 0.60 Shoulder dystocia Metformin group = 6 (1.7%) Insulin group = 11 (3.0%) p = 0.33 Fetal death Metformin Group = 0/363 Insulin Group = 1/370	
Full citation Silva,J.C., Fachin,D.R., Coral,M.L., Bertini,A.M., Perinatal impact of the use of metformin and glyburide for the	Sample size N = 200 women diagnosed with gestational diabetes using WHO criteria who attended one of 3 hospitals in Joinville, Brazil. Women were screened using home glucose self monitoring by capillary glucose testing 7 days after initial instruction, assessing fasting and postprandial values. Acceptable values 90mg/dl and postprandial 120mg/dl, Women	Interventions Women were randomised to treatment between 11 and 33 gestational weeks. Glibenclamide: An initial dose of	Details Diet: All women were given instructions for a diet designed to provide 35kcals/kg at normal body weight and 25kcals/kg at obese body weight, with 35-45% calories from	Results Maternal outcomes Treatment failure (need to change therapy to insulin) Glibenclamide group = 28/96 Metformin group = 22/104 p=0.56 Neonatal outcomes Fetal hypoglycaemia	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials Appropriate randomisation

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
treatment of gestational diabetes mellitus, Journal of Perinatal Medicine, 40, 225-228, 2012 Ref Id 177659 Country/ies where the study was carried out Brazil Study type Randomised controlled trial Aim of the study To evaluate the perinatal impact of metformin and glibenclamide in the treatment of gestational diabetes mellitus Study dates 1 July 2008 to 30 September 2010 Source of funding Not stated, but the researchers had no link or interests with the manufacturers of the drugs or	were offered participation in the study if 2 values were abnormal. Glibenclamide group =96 Metformin group = 104 Two women with intrauterine death were excluded (one in each group). Characteristics Inclusion criteria Inclusion criteria were minimum age 18 years, gestational age 11-33 weeks, single gestation, fetal abdominal circumference within normal percentile (>10% and <75%)	2.5mg before breakfast and dinner was increased as necessary by 2.5 - 5mg weekly until glucose control was acheived or until a maximum dose of 20mg/day was reached. Metformin: An initial dose of 500mg before breakfast and dinner was increased as necessary by 500-1000 mg weekly until glucose control was acheived or until a maximum dose a maximum dose of 2500 mg/day was reached.	carbohydrates and consisting of 3 full meals and four light meals. Exercise: No details are given regarding the rexrcise regimen woem were to follow Monitoring: All women performed home glucose self monitoring of fasting and postprandial capillary glucose testing to adjust dosage of medication. Insulin therapy was started at 0.7 IU/kg/day regular insulin preprandial and neutral protamine hagedorn (NHP) insulin at bedtime when glycaemic goals were not met. Statistical analysis Variables were analysed descriptively using calculations of means, standard deviations, absolute and relative frequencies. Student's t-tests and Mann-Whitney U tests were used to test the equality of the means of the two groups. Fisher's exact tests and the X2 tests were used to test group homogeneity for categorical variables. The	(<40mg/dl) Glibenclamide group = 13/96 Metformin group = 11/104 p=0.81 Large for gestational age (percentile above 90 in growth curves) Glibenclamide group = 19/96 Metformin group = 9/104 p=0.08 NICU admission (no definition given) Glibenclamide group = 7/96 Metformin group = 9/104 p=0.94 Death (no further definition given) Glibenclamide group = 1/96 Metformin group = 1/104 p=0.99	method: Yes Adequate allocation concealment: Yes Groups comparable at baseline: No not for all characteristics. Women in the glibenclamide group on average were heavier and had had fewer babies previously Groups received the same care (apart from the intervention): Yes Participants kept 'blind' to allocation: No - It was an open RCT Care givers kept 'blind' to allocation: No - It was an open RCT Follow up equal for groups: Yes How many participants did not complete treatment in each group?: None Were the groups were comparable for treatment completion: Yes For how many participants in each group were no outcome data

Characteristics

	Glibenclamide (n=96)	Metformin (n=104)	p value
Age (yrs)	31.29±5.36	32.63±5.61	0.09
Gestations	2.47±1.30	2.84±1.25	0.04
GA (wks)at inclusion	25.44±7.13	26.96±6.44	0.11
BMI	28.61±5.88	28.69±5.37	0.46
Weight gain (kg)	9.84±6.42	7.78±7.42	0.04
OGTT fasting (mg/dl)	94.04±16.25	95.84±20.91	0.52
OGTT 2h (mg/dl)	160.83±18.60	165.59±21.80	0.12

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td>equipment reported in the study</td><td>and absence of other pathologies that might interfere with perinatal results or hypoglycaemic therapy.

Exclusion criteria
Exclusion criteria were intolerance of the drugs or unwillingness to participate, fetal risk (fetal abdominal circumference at percentile >97% or <5%), lack of follow up or fetal malformation diagnosed on delivery.</td><td></td><td>significance level of the tests was 0.05.</td><td></td><td>available?: None
The groups were comparable with respect to the availability of outcome data: Yes
Appropriate length of follow-up: Yes
Precise outcome definitions used: No not for all outcomes
Outcome determined using valid and reliable methods: Yes
Investigators kept 'blind' to allocation: Unclear
Investigators kept 'blind' to other important confounding and prognostic factors: Unclear

Other information
None.</td></tr>
<tr><td>Full citation
Spaulonci,C.P., Bernardes,L.S., Trindade,T.C., Zugaib,M., Francisco,R.P., Randomized trial of metformin vs insulin in the management of gestational</td><td>Sample size
N = 94

Characteristics<table><tr><th>Characteristic</th><th>Metformin</th><th>Insulin</th><th>P-value</th></tr><tr><td>Mean age, years</td><td>31.93 ± 6.02</td><td>32.76 ± 4.66</td><td>0.46</td></tr></table></td><td>Interventions
Intervention
Metformin

Control
Insulin</td><td>Details
Eligible women who met inclusion criteria were randomly assigned to receive either metformin (n = 46) or insulin (n = 46). Two women (one from each group) were excluded.</td><td>Results
Caesarean delivery
Metformin: 33/46
Insulin: 30/46
RR = 1.10 (95% CI 0.83 to 1.45)*

Neonatal hypoglycaemia
Metformin: 3/46</td><td>Limitations
NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual
A. Selection bias
A1: An appropriate</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
diabetes, American Journal of Obstetrics and Gynecology, 209, 34-37, 2013 Ref Id 305716 Country/ies where the study was carried out Brazil Study type Randomised controlled trial. Aim of the study To compare glycaemic control in women who take metformin versus insulin for the treatment of GDM and to identify predictors of the need for insulin in women initially treated with metformin. Study dates November 1st 2007 to January 31st 2010. Source of funding	<table><tr><td>Median number of pregnancies (IQR)</td><td>2 (1 to 8)</td><td>3 (1 to 8)</td><td>0.04</td></tr><tr><td>Median parity (IQR)</td><td>1 (0 to 5)</td><td>1 (0 to 6)</td><td>0.72</td></tr><tr><td>Mean gestational age at diagnosis, weeks</td><td>30.40 ± 3.71</td><td>30.63 ± 3.35</td><td>0.76</td></tr><tr><td>Mean BMI at diagnosis, kg/m2</td><td>31.97 ± 4.71</td><td>31.31 ± 5.80</td><td>0.55</td></tr><tr><td>Mean HbA_{1c} at diagnosis, %</td><td>5.90 ± 0.75</td><td>5.93 ± 0.80</td><td>0.86</td></tr></table> Inclusion criteria Singleton pregnancies Use of diet and exercise for at least one week without obtaining glycaemic control Absence of risk factors for lactic acidosis Absence of anatomical or chromosomal fetal abnormalities detected by ultrasound Exclusion criteria Not reported.		Unsatisfactory glycaemic control was defined as > 30% of capillary blood glucose values above reference values 1 week after the initiation of diet therapy and physical activity. Glucose reference values were not reported. Outcomes included: Caesarean delivery Neonatal hypoglycaemia (not defined) Macrosomia (not defined) Statistical analysis Logistic regression was used to identify predictors of the need for supplemental insulin therapy in women treated with metformin.	Insulin: 10/46 RR = 0.30 (95% CI 0.09 to 1.02)* Macrosomia Metformin: 0/46 Insulin: 3/46 RR = 0.14 (95% CI 0.007 to 2.64)* Treatment failure Metformin: 12/46 Insulin: not reported RR not calculable *Calculated by the NCC-WCH technical team.	method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Unclear - method of randomisation was not described. A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Unclear A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Unclear B. Performance bias B1: The comparison groups received the same care apart from the intervention(s)

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Not reported.					studied. Unclear B2: Participants receiving care were kept 'blind' to treatment allocation. No B3: Individuals administering care were kept 'blind' to treatment allocation. Unclear C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? Not reported. b. The groups were comparable for treatment completion (that is, there were no important or systematic differences

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					between groups in terms of those who did not complete treatment). Unclear C3: a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes D2: The study used a precise definition of outcome. No - no outcomes were defined.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					D3: A valid and reliable method was used to determine the outcome. Unclear D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information Minimal baseline characteristics were reported by the study and methodology was not fully described.
Full citation Tertti,K., Ekblad,U., Koskinen,P., Vahlberg,T., Ronnemaa,T., Metformin vs. insulin in gestational	Sample size N = 221 Characteristics Characteristic / Metformin / Insulin / P-value Mean maternal / 31.9 ± 5.0 / 32.1 ± / 0.80	Interventions Intervention Metformin was initiated at a dose of 500mg once daily for the first two days, increased to twice	Details The Finnish national criteria for diagnosing GDM changed during the study. Consequently OGTT tests were performed if one or more of the following were present:	Results Large for gestational age (< 90th percentile) Metformin: 16/109 Insulin: 17/107 RR = 0.92 (95% CI 0.49 to 1.72)*	Limitations NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
diabetes. A randomized study characterizing metformin patients needing additional insulin, Diabetes, Obesity and Metabolism, 15, 246-251, 2013 Ref Id 248278 Country/ies where the study was carried out Finland Study type Randomised controlled trial. Aim of the study To assess whether metformin is as effective as insulin in treating women with gestational diabetes mellitis with respect to fetal weight gain. The study also aimed to identify predictors of the need for insulin therapy in women treated with metformin.	<table><tr><td>age, years</td><td></td><td>5.4</td><td></td></tr><tr><td>Primipara, n (%)</td><td>42 (38.2)</td><td>48 (44.9)</td><td>0.45</td></tr><tr><td>Mean BMI, kg/m2</td><td>29.4 ± 5.9</td><td>28.9 ± 4.7</td><td>0.74</td></tr><tr><td>Mean gestational age at randomisation, weeks</td><td>30.3 ± 2.0</td><td>30.4 ± 1.8</td><td>0.72</td></tr></table> Inclusion criteria Singleton pregnancies Presence of GDM diagnosed based on two or three abnormal 2 hour plasma glucose values from a 75g OGTT Met criteria to start medication for GDM Exclusion criteria Cardiac or renal insufficiency Liver disease Metformin use within three months preceding pregnancy or during pregnancy before the OGTT Self-measured plasma glucose values > 7.0mmol/l or 1 hour postprandial values > 11.0mmol/l	daily for the first week. The dose was increased to a maximum of 1g twice daily if required. Target values were < 5.5mmol/l after an overnight fast and < 7.8mmol/l 1 hour postprandial. Insulin was added if these targets were not met with metformin alone. Control Insulin treatment comprised NPH insulin and/or rapid acting insulin lispro or aspart.	BMI ≥ 25kg/m2 Aged ≥ 40 years Previous macrosomic child Suspected fetal macrosomia in the current pregnancy Glucosuria Weight gain ≥ 20kg during pregnancy GDM in a previous pregnancy Diagnostic cut-offs until December 2008 were: Fasting blood glucose ≥ 4.8mmol/l 1 hour postprandial ≥ 10.0mmol/l 2 hour postprandial ≥ 8.7mmol/l After 2008 cut-offs were as follows: Fasting blood glucose ≥ 5.3mmol/l 1 hour postprandial ≥ 10.0mmol/l 2 hour postprandial ≥ 8.6mmol/l All women attended the hospital for dietary counselling and were taught to measure overnight fasting and 1 hour postprandial glucose at least four	NICU stay Metformin: 34/109 Insulin: 39/107 RR = 0.86 (95% CI 0.59 to 1.25)* Neonatal hypoglycaemia Metformin: 18/109 Insulin: 18/107 RR = 0.98 (95% CI 0.54 to 1.78)* Caesarean section Metformin: 15/109 Insulin: 18/107 RR = 0.82 (95% CI 0.44 to 1.54)* Induction of labour Metformin: 42/109 Insulin: 58/107 RR = 0.71 (95% CI 0.53 to 0.95)* Assisted vaginal delivery Metformin: 9/109 Insulin: 8/107 RR = 1.10 (95% CI 0.44 to 2.74)* Treatment failure Metformin: 23/110 Insulin: not reported RR not calculable	A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Unclear - sealed envelopes were used but the method of randomisation was not described. A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). No A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes B. Performance bias B1: The

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Study dates June 2006 to December 2010. Source of funding Grants from the Finnish Diabetes Association and EVO (grant number 3857).			times daily. Criteria for pharmacological treatment were: Two or more fasting blood glucose values ≥ 5.5mmol/l, and/or Postprandial values ≥ 7.8mmol/l Women were randomised between 22 and 34 weeks' gestation (metformin = 111, insulin = 110) using sealed envelopes. Clinical appointments were every one to two weeks throughout the remainder of the pregnancy. Large for gestational age was defined as birth weights > 2 standard deviations above the mean (approximately 97.5th percentile). Data were also provided for birth weights > 90th percentile. NICU stay was not defined. Neonatal hypoglycaemia was defined as <	*Calculated by the NCC-WCH technical team.	comparison groups received the same care apart from the intervention(s) studied. Yes B2: Participants receiving care were kept 'blind' to treatment allocation. No B3: Individuals administering care were kept 'blind' to treatment allocation. No C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Yes C2: a. How many participants did not complete treatment in each group? One in the metformin group and three in the insulin group. b. The groups were comparable for

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			2.6mmol/l and requiring intravenous glucose treatment. Three women in the insulin group were excluded due to refusal to start insulin after randomisation. One woman in the metformin group was excluded as she moved away from the local area during the study. Statistical analysis Continuous variables were compared between groups using either the Mann-Whitney U test or two-sample t-test. Poisson regression was used to analyse dichotomous variables between groups. Relative risks and 95% confidence intervals were calculated. P-values < 0.05 were considered to be statistically significant.		treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Yes C3: a. For how many participants in each group were no outcome data available? Unclear b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					D2: The study used a precise definition of outcome. No - NICU stay was not defined. D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information None
Full citation Thompson,D.J., Porter,K.B., Gunnells,D.J., Wagner,P.C., Spinnato,J.A., Prophylactic insulin in the	Sample size Total sample size comprised 108 women (68 successfully completed treatment: 34 intervention, 34 control). Characteristics	Interventions Intervention Diet plus 20 units of NPH insulin and 10 units of regular insulin 30 mins before breakfast.	Details All consenting women who attended for prenatal care at the University of South Alabama Medical Center were studied. Patients were screened at 28 weeks. Those who	Results Caesarean (includes those who failed treatment) Diet + insulin: 14/45 Diet alone: 16/50 RR = 0.97 (95% CI 0.54 to 1.76)*	Limitations NICE checklist for randomised controlled trials, taken from Appendix C of the NICE guidelines manual

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
management of gestational diabetes, Obstetrics and Gynecology, 75, 960-964, 1990 Ref Id 177702 Country/ies where the study was carried out United States of America Study type Randomised controlled trial. Aim of the study To determine whether insulin plus diet reduces maternal and neonatal morbidity compared with diet alone in women with gestational diabetes mellitus. Study dates October 1985 to June 1988. Source of	<table><tr><td>Characteristic</td><td>Diet</td><td>Diet+insulin</td></tr><tr><td>Mean age, years</td><td>26 ± 5.7</td><td>27 ± 5.4</td></tr><tr><td>Gravidity</td><td>2.5 ± 1.5</td><td>3.0 ± 1.7</td></tr><tr><td>Parity</td><td>1.3 ± 1.4</td><td>1.4 ± 1.5</td></tr><tr><td>Weight at 20 weeks, lb</td><td>184 ± 46</td><td>175 ± 38</td></tr><tr><td>3 hour OGTT fasting glucose</td><td>101 ± 16</td><td>101 ± 26</td></tr></table> All between group comparisons were non-significant. P-values were not reported. Inclusion criteria Women with gestational diabetes who consented to be enrolled in the trial. Following a 50g fasting and 1 hour glucose screening test at 28 weeks' gestation, women with fasting values ≥ 105mg/dl or a 1 hour value ≥ 140mg/dl were referred for a 3 hour OGTT. OGTT cut-offs for inclusion in the study were > 105mg/dl fasting, > 190mg/dl at 1 hour, > 165mg/dl at 2 hours and > 145mg/dl at 3 hours. GDM was diagnosed if any two values were abnormal. Exclusion criteria	Control 35kcal/kg ideal body weight/day comprising 50% kcal as carbohydrate, 30% as fat and 20% as protein.	met screening criteria were referred for a 3 hour OGTT. Following diagnosis with gestational diabetes women were allocated to either a standard diet group or diet plus insulin. Allocation was random using sealed envelopes. Subjects were considered to have failed treatment if fasting glucose levels > 105mg/dl once or 2 hour post-prandial levels > 120mg/dl twice. Failed subjects in the diet group had insulin added; those in the insulin group had higher insulin doses. Successes were those who maintained glycaemic control; no self-monitoring of blood glucose was performed. All undelivered pregnancies were induced at 42 weeks. Outcomes included: Perinatal mortality Perinatal morbidity (birth trauma) Macrosomia (> 4000g) Hypoglycaemia (plasma	Treatment failure Diet + insulin: 9/45 Diet alone: 16/50 RR = 0.63 (95% CI 0.04 to 9.90)* Macrosomia (successes only) Diet + insulin: 2/34 Diet alone: 9/34 RR = 0.20 (95% CI 0.05 to 0.86)* Hypoglycaemia (successes only) Diet + insulin: 2/34 Diet alone: 5/34 RR = 0.40 (95% CI 0.08 to 1.92)* Perinatal mortality (successes only) Diet + insulin: 0/34 Diet alone: 0/34 RR not calculable. Shoulder dystocia (successes only) Diet + insulin: 0/34 Diet alone: 0/34 RR not calculable. *Calculated by the NCC-WCH technical team.	A. Selection bias A1: An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups). Yes A2: There was adequate concealment of allocation (such that investigators, clinicians and participants cannot influence enrolment or treatment allocation). Yes A3: The groups were comparable at baseline, including all major confounding and prognostic factors. Yes B. Performance bias B1: The comparison groups received the same care apart from the intervention(s) studied. Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
funding Not reported.	Patients with a fasting 3 hour OGTT measurement ≥ 140mg/dl, those who refused to participate, were diagnosed after 36 weeks' gestation or participated for less than 6 weeks.		glucose < 30mg/dl) Statistical analysis Categorical data were analysed using Yates corrected X2 tests. Comparisons of group means were made using two-tailed t-tests for independent samples. Results were considered significant for p-values < 0.05.		B2: Participants receiving care were kept 'blind' to treatment allocation. Yes B3: Individuals administering care were kept 'blind' to treatment allocation. Yes C. Attrition bias C1: All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up). Unclear C2: a. How many participants did not complete treatment in each group? Treatment failures: 9 in the diet + insulin group, 16 in the diet alone group. Not clear if these women did not complete treatment. b. The groups were comparable for

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment). Unclear C3: a. For how many participants in each group were no outcome data available? Not reported. b. The groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available). Unclear D. Detection bias D1: The study had an appropriate length of follow-up. Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					D2: The study used a precise definition of outcome. Yes D3: A valid and reliable method was used to determine the outcome. Yes D4: Investigators were kept 'blind' to participants' exposure to the intervention. Unclear D5: Investigators were kept 'blind' to other important confounding and prognostic factors. Unclear Other information None.

A.16 Timing of Birth

What is the gestational age-specific risk of intrauterine death in pregnancies with type 1, type 2 or gestational diabetes, and the optimal timing of birth?

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td>Full citation
Rosenstein,M.G., Cheng,Y.W., Snowden,J.M., Nicholson,J.M., Doss,A.E., Caughey,A.B., The risk of stillbirth and infant death stratified by gestational age in women with gestational diabetes, American Journal of Obstetrics and Gynecology, 206, 309-7, 2012
Ref Id
236324
Country/ies where the study was carried out
USA
Study type
Retrospective cohort study

Aim of the study</td>
<td>Sample size
4,190,953 non-anomalous singleton deliveries with gestational ages between 36 weeks and 0 days and 42 weeks 6 days.

Characteristics
<table>
<tr><td></td><td>Women with gestational diabetes

N=193,028</td><td>Women without gestational diabetes N=3,997,925</td><td>p - value</td></tr>
<tr><td>Maternal Age (years: mean ± SD)</td><td>31.4 ± 5.8</td><td>27.7 ± 6.2</td><td><0.001</td></tr>
<tr><td>Ethnicity</td><td></td><td></td><td><0.001</td></tr>
<tr><td>White N (%)</td><td>52,498 (27.2%)</td><td>1,504,878 (37.7%)</td><td></td></tr>
<tr><td>African-American N (%)</td><td>7,548 (3.9%)</td><td>217,883 (5.5%)</td><td></td></tr>
<tr><td>Latino N</td><td>94,682</td><td>1,766,57</td><td></td></tr>
</table>
</td>
<td>Interventions
Delivery at a given week of gestation was compared to expectant management (continuation of the pregnancy for another week with delivery one week later)</td>
<td>Details
Mortality risk of delivery at a given week (Definition: the rate among those neonates born at that week of gestation) was compared with a composite mortality risk of a week of expectant management (Definition: the risk of stillbirth over that week plus the mortality risk experienced by infants born in the subsequent week of gestation) at different gestational ages among women with gestational diabetes.

Infant mortality (Definition: age 29 – 365 days of life) was examined rather than neonatal death</td>
<td>Results
Incidence of stillbirth
<table>
<tr><td></td><td>GDM</td><td>GDM</td><td>No GDM</td><td>No GDM</td><td></td></tr>
<tr><td>GA</td><td>Total Deliveries</td><td>Stillbirth/ 1000 deliveries (95% CI)</td><td>Total Deliveries</td><td>Stillbirth/ 1000 deliveries (95% CI)</td><td>RR Stillbi (95% CI)</td></tr>
<tr><td>36</td><td>10445</td><td>6.13</td><td>155597</td><td>5.43</td><td>1.13 (0.88 1.45)</td></tr>
<tr><td>37</td><td>22157</td><td>3.38</td><td>340239</td><td>1.34</td><td>1.34 (1.06 1.70)</td></tr>
<tr><td>38</td><td>44487</td><td>1.51</td><td>736413</td><td>1.37</td><td>1.10 (0.86 1.41)</td></tr>
<tr><td>39</td><td>56085</td><td>1.18</td><td>1105279</td><td>0.91</td><td>1.30 (1.01 1.66)</td></tr>
<tr><td>40</td><td>37819</td><td>0.90</td><td>981106</td><td>0.74</td><td>1.21 (0.86 1.71)</td></tr>
<tr><td>41</td><td>15739</td><td>1.21</td><td>510292</td><td>0.85</td><td>1.42 (0.90</td></tr>
</table>
</td>
<td>Limitations
NICE guidelines manual. Appendix I: Methodology checklist: Prognostic studies

1) The study sample represents the population of interest with regard to key characteristics, sufficient to limit potential bias to the results: N, the</td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td>To compare the stillbirth and infant mortality risks between delivery and expectant management in women with gestational diabetes

Study dates
1997 to 2006

Source of funding
One author was supported by the National Institute of Child Health and Human Development</td>
<td>
<table>
<tr><td>(%)</td><td>(49.1%)</td><td>9 (44.2%)</td><td></td></tr>
<tr><td>Asian N (%)</td><td>35,295 (18.3%)</td><td>443,980 (11.1%)</td><td></td></tr>
<tr><td>Other N (%)</td><td>2,877 (1.5%)</td><td>59,816 (1.5%)</td><td></td></tr>
<tr><td>Preeclampsia N (%)</td><td>7,827 (4.1%)</td><td>84,588 (2.1%)</td><td><0 01</td></tr>
<tr><td>Chronic Hypertension N (%)</td><td>4,574 (2.4%)</td><td>22,325 (0.6%)</td><td><0 01</td></tr>
<tr><td>Gestational age at delivery (weeks: mean, SD)</td><td>38.8 ± 1.4</td><td>39.1 ± 1.4</td><td><0 01</td></tr>
<tr><td>Birthweight (grams: mean, SD)</td><td>3,475 ± 541</td><td>3,415 ± 475</td><td><0 01</td></tr>
<tr><td>Education (≥12 years) N (%)</td><td>71,014 (43.5%)</td><td>1,496,734 (42.6%)</td><td><0 01</td></tr>
</table>

Inclusion criteria
Women were identified from California Vital Statistics Birth Certificate Data linked with the California Patient Discharge Data as well as Vital Statistics Death
</td>
<td></td>
<td>(Definition: death within 28 days of birth)
because previous data demonstrated that term infants who died within the first year of life were more likely to do so in the post-neonatal period than in the neonatal period and because of its significant magnitude and association with gestational age at delivery.
Incidence of stillbirth at a given gestational age was defined as the number of stillbirths at that gestational age per 1000 deliveries.

Infant mortality at each gestational age was defined as the number of infants born at this gestational age who die within one year of life per 10,000 live births at that same</td>
<td>
<table>
<tr><td></td><td></td><td></td><td></td><td></td><td>2.25)</td></tr>
<tr><td>42</td><td>6296</td><td>0.95</td><td>168,999</td><td>1.15</td><td>0.83 (0.37 1.86)</td></tr>
</table>

Incidence of Neonatal death
<table>
<tr><td></td><td>GDM</td><td>GDM</td><td>No GDM</td><td>No GDM</td><td></td></tr>
<tr><td>GA</td><td>Deliveries</td><td>Neonatal death/ 10,000 live births (95% CI)</td><td>Deliveries</td><td>Neonatal death/ 10,000 live births (95% CI)</td><td>RR (95% CI) Neonatal deat</td></tr>
<tr><td>36</td><td>10,375</td><td>10.6 (5.3 - 19.0)</td><td>154579</td><td>9.1 (7.7 - 10.8)</td><td>1.16 (0.63 2.14</td></tr>
<tr><td>37</td><td>22,074</td><td>6.8 (3.8 - 11.2)</td><td>339187</td><td>6.1 (5.3 - 7.0)</td><td>1.11 (0.66 1.88</td></tr>
<tr><td>38</td><td>44,414</td><td>3.6 (2.1 - 5.9)</td><td>735205</td><td>3.9 (3.5 - 4.4)</td><td>0.92 (0.56 1.53</td></tr>
<tr><td>39</td><td>56,011</td><td>3.4 (2.0 - 5.3)</td><td>1104127</td><td>2.8 (2.5 - 3.1)</td><td>1.21 (0.76 1.92</td></tr>
<tr><td>40</td><td>37,779</td><td>2.6 (1.3 - 4.9)</td><td>980203</td><td>3.4 (3.1 - 3.8)</td><td>0.78 (0.41 1.46</td></tr>
</table>
</td>
<td>largest ethnic group is Latin American which is not directly applicable to the UK
2) Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias: Y
3) The prognostic factor of interest is adequately measured in study participant</td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td></td>
<td>Certificate Data and Vital Statistics Fetal Death File. 193,028 deliveries were to women with a diagnosis of gestational diabetes identified from maternal medical records using ICD-9 codes: 648.8, 648.80, 648.81, 648.82, 648.83 and 648.84.

Exclusion criteria
Women with a diagnosis of pre-pregnancy (Type 1 or Type 2) diabetes mellitus were excluded (using ICD-9 codes: 648.0, 648.01, 648.02, 648.03, and 648.04). Multiple gestations and births with congenital anomalies as determined by diagnosis codes on the birth certificate and the infant's medical record (ICD-9 codes Q00-Q99) were also excluded.

The mother/infant pair was excluded from analysis if the date of last menstrual period was missing or was nonsensical, as this was needed to calculate the length of gestation</td>
<td></td>
<td>gestational age.

Calculations relied on the following assumptions:
1. The risk of infant death has a uniform distribution throughout the week of gestation.
2. When estimating the risk of delivering at a particular gestational age, the fetus is not at risk for stillbirth beyond that gestational age, therefore their mortality risk in that week is equal only to the risk of infant death.
3. The composite risk associated with expectant management is the sum of the risk of stillbirth during the week of gestation plus the risk of infant death in the following week of gestation.</td>
<td>
<table>
<tr><td>41</td><td>15,717</td><td>3.2 (1.0 - 7.4)</td><td>509749</td><td>3.6 (3.1 - 4.2)</td><td>0.88 (0.36 2.14</td></tr>
<tr><td>42</td><td>6,285</td><td>6.4 (1.7 - 16.3)</td><td>168769</td><td>4.7 (3.7 - 5.8)</td><td>1.36 (0.5(3.72</td></tr>
</table>
<table>
<tr><td></td><td>GDM</td><td>GDM</td><td>No GDM</td><td>No GDM</td><td></td></tr>
<tr><td>GA</td><td>Deliveries</td><td>Infant death/ 10,000 live births (95% CI)</td><td>Deliveries</td><td>Infant death/ 10,000 live births (95% CI)</td><td>RR (95 CI) of In nt dea</td></tr>
<tr><td>36</td><td>10445</td><td>19.3 (11.8 - 29.8)</td><td>155597</td><td>22.9 (20.6 - 25.4)</td><td>0.8 (0.5 1.32</td></tr>
<tr><td>37</td><td>22,157</td><td>14.0 (9.5 - 19.9)</td><td>340,239</td><td>18.4 (17.0 - 19.9)</td><td>0.7 (0.5 1.1)</td></tr>
<tr><td>38</td><td>44,487</td><td>10.6 (7.8 - 14.1)</td><td>736,413</td><td>13.3 (12.5 - 14.2)</td><td>0.8 (0.5 1.06</td></tr>
<tr><td>39</td><td>56,085</td><td>8.7 (6.5 - 13.2)</td><td>1,105,279</td><td>10.7 (10.1 - 11.4)</td><td>0.8 (0.6 1.08</td></tr>
<tr><td>40</td><td>37,819</td><td>9.5 (6.7 - 13.2)</td><td>981,106</td><td>11.6 (10.9 - 12.3)</td><td>0.8 (0.5 1.14</td></tr>
</table>
</td>
<td>s, sufficient to limit potential bias: Y
4) The outcome of interest is adequately measured in study participants, sufficient to limit potential bias: Y
5) Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest:</td>
</tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				41 | 15,739 | 11.5 (6.8 - 18.1) | 510,292 | 12.8 (11.9 - 13.9) | 0.8 (0.5 1.4) 42 | 6,296 | 9.5 (3.5 - 20.8) | 168,999 | 14.0 (12.3 - 15.9) | 0.6 (0.3 1.5) * Calculated by NCC-WCH	N, the groups were significantly different at baseline for key characteristics, most relevantly women with gestational diabetes were significantly more likely to have hypertensive disorders than those without gestational diabetes 6) The statistical analysis is appropriate for the design of the study, limiting potential

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>for the presentation of invalid results: Y

Other information</td></tr>
<tr><td>Full citation
Holman,N., Bell,R., Murphy,H., Maresh,M., Women with pre-gestational diabetes have a higher risk of stillbirth at all gestations after 32 weeks, Diabetic MedicineDiabet. Med., n/a-n/a, 2014
Ref Id
319500
Country/ies where the study was carried out
England
Study type
Retrospective analysis of audit</td><td>Sample size
Data on stillbirth from pregnant women with diabetes prior to pregnancy (n=2085) were compared with stillbirth data for all births in England and Wales for 2007, 2008, 2010 and 2011 (n=3,522,869) obtained from the Office of National Statistics.

Characteristics
Of 2085 women with diabetes prior to pregnancy: 1154 (55.8%) Type 1 diabetes and 895 (43.7%) Type 2 diabetes.

Inclusion criteria
Singleton pregnancy

Exclusion criteria
Births associated with major congenital malformations</td><td>Interventions
Not relevant</td><td>The number of live births and stillbirths by gestation were identified. Stillbirth was defined as an infant born after 24 completed weeks of gestation that did not show any signs of life after birth.

The stillbirth rate was calculated using the number of stillbirths at a specific gestational age divided by the total births (live and still) at that specific gestational age</td><td>Results<table><tr><td></td><td>Type 1&2 diabetes</td><td>Type 1&2 diabetes</td><td>All births E&W</td><td>All births E&W</td><td></td></tr><tr><td>GA (wks)</td><td>Total deliveries</td><td>Stillbirth/ 1000 total births (95% CI)</td><td>Total deliveries</td><td>Stillbirth/ 1000 total births (95% CI)</td><td>RR (95%)</td></tr><tr><td>24-27</td><td>20</td><td>250 (89.8-490.8)</td><td>16927</td><td>264 (257.2 – 272.6)</td><td>0.95 (0.82 - 1.10)</td></tr><tr><td>28-31</td><td>49</td><td>81.6 (29.5 – 194.6)</td><td>31894</td><td>93.5 (90.2 – 96.9)</td><td>0.87 (0.66 - 1.16)</td></tr><tr><td>32-34</td><td>161</td><td>43.5 (20.6 – 87.7)</td><td>69930</td><td>34.8 (33.5 – 36.2)</td><td>1.25 (0.81 - 1.94</td></tr></table></td><td>Limitations
NICE guidelines manual. Appendix I: Methodology checklist: Prognostic studies

1) The study sample represents the population of interest with regard to key characteristics, sufficient</td></tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td>data

Aim of the study
To explore the additional risk of stillbirths and to quantify that risk according to gestational age among women with diabetes

Study dates
Audit data on pregnancies of women with pre-gestational diabetes from two cohorts: from 3 regions (Northern, North West and East Anglia) in 2007 and 2008 and from 1 region (East Anglia) and from 13 other units in England in 2010 and 2011.

Source of funding
None stated</td><td></td><td></td><td></td><td><table>
<tr><td></td><td></td><td></td><td></td><td></td><td>)</td></tr>
<tr><td>35-36</td><td>392</td><td>10.2 (3.9 – 26.0)</td><td>143609</td><td>13.6 (13.0 – 14.2)</td><td>0.75 (0.33 - 1.68)</td></tr>
<tr><td>37-38</td><td>1185</td><td>5.1 (2.3 – 11.0)</td><td>670426</td><td>3.5 (3.3 – 3.6)</td><td>1.46 (0.37 - 5.66)</td></tr>
<tr><td>≥39</td><td>278</td><td>10.8 (3.6 – 31.3)</td><td>2590083</td><td>1.5 (1.4 – 1.5)</td><td>7.2 (1.31 - 39.63)</td></tr>
</table></td><td>to limit potential bias to the results: Yes
2) Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias: Yes
3) The prognostic factor of interest is adequately measured in study participants, sufficient to limit potential</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					bias: Yes 4) The outcome of interest is adequately measured in study participants, sufficient to limit potential bias: Yes 5) Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest: Yes 6) The statistical analysis is

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>appropriate for the design of the study, limiting potential for the presentation of invalid results: Yes

Other information</td></tr>
<tr><td>Full citation
Eidem,I., Vangen,S., Hanssen,K.F., Vollset,S.E., Henriksen,T., Joner,G., Stene,L.C., Perinatal and infant mortality in term and preterm births among women with type 1 diabetes, Diabetologia, 54, 2771-2778, 2011
Ref Id
236459</td><td>Sample size
Record linkage of two nationwide registries allowed identification of 1,307 babies born to women with pregestational type 1 diabetes and 1,161,092 births in the background population to mothers without type 1 diabetes.

Characteristics
<table><tr><th></th><th>Type 1 diabetes (n=1307)</th><th>Background population (n=1,161,092)</th></tr></table></td><td>Interventions
Perinatal mortality rates by gestational age were calculated. Perinatal death was defined as stillbirth (death of the foetus before or during labour) or early</td><td>Details
Record linkage of two nationwide registries allowed identification of babies born to women with pregestational type 1 diabetes (Norwegian Childhood Diabetes Registry) and births in the background population to mothers without type 1 diabetes (Medical Birth Registry of Norway) during the</td><td>Results
<table><tr><th></th><th>Type 1 diabetes</th><th>Type 1 diabetes</th><th>No type 1 diabetes</th><th>No type 1 diabetes</th><th></th></tr><tr><td>GA</td><td>Total deliveries</td><td>Perinatal mortality/ 1000 deliveries (95% CI)</td><td>Total deliveries</td><td>Perinatal mortality/ 1000 deliveries (95% CI)</td><td>RR (95% CI)</td></tr><tr><td>32-34</td><td>85</td><td>58.8 (19.4-132.0)</td><td>19,594</td><td>50.3 (47.3-53.5)</td><td>1.17 (0.50 2.74</td></tr><tr><td>35-36</td><td>190</td><td>15.8 (3.27-</td><td>39,553</td><td>19.0 (17.7-</td><td>0.83 (0.27</td></tr></table></td><td>Limitations
NICE guidelines manual. Appendix I: Methodology checklist: Prognostic studies

1) The study sample represents the</td></tr>
</table>

Study details	Participants			Interventions	Methods	Outcomes and Results						Comments
Country/ies where the study was carried out Norway Study type Retrospective cohort study Aim of the study To estimate the risks of adverse birth outcomes (eg stillbirth, infant death, preterm birth and pre-eclampsia) in women with type 1 diabetes, compared with the background population. Study dates Data held in the registry between the years 1985 to 2004 was investigated. Source of funding Supported by research grants from the South-Eastern Norway	Age at diagnosis of diabetes (years) Median (IQ range)	11 (8-13)	-	neonatal death (death during the first 7 days of life).	period 1985–2004. Logistic regression was used to estimate the relative risks of birth outcomes in pregnancies with type 1 diabetes compared with the background population before and after adjusting for confounding factors. Perinatal mortality was plotted by gestational age for the two groups			45.5)		20.4)	2.56	population of interest with regard to key characteristics, sufficient to limit potential bias to the results: Yes 2) Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias: Yes 3) The prognostic factor of interest is adequatel
	Duration of diabetes (years) Median (IQ range)	17 (12-21)	-			37	152	13.2 (1.60-46.7)	47,517	9.28 (8.44-10.2)	1.42 (0.36 5.63	
	Age at delivery (years) Median (IQ range)	27 (24-30)	28 (25-32)			38	225	8.89 (1.08-31.7)	105,234	4.51 (4.12-4.94)	1.97 (0.49 7.85	
	Parity (%)					39	245	12.2 (2.53-35.4)	206,321	2.88 (2.66-3.12)	4.25 (1.38 13.1	
	Para 0	50.2	41.6			40	159	6.29 (0.16-34.5)	281,805	2.08 (1.91-2.25)	3.03 (0.43 21.4	
	Para 1	34.5	35.3			41-45	1071	29.7 (6.17-84.4)	366,653	2.39 (2.24-2.56)	12.4 (4.06 37.9	
	Para 2	12.4	16.6									
	Para 3	2.3	4.5									
	Para 4 or more	0.6	1.9									
	Educational level (%)											
	<12 years of school	34.7	35.0									

Study details	Participants			Interventions	Methods	Outcomes and Results	Comments
Regional Health Authority, Oslo Diabetes Research Centre and the Norwegian Research Council (for initiation of the study)	Completed 12 years of school	32.1	29.7				y measured in study participants, sufficient to limit potential bias: Yes 4) The outcome of interest is adequately measured in study participants, sufficient to limit potential bias: Yes 5) Important potential confounders are appropriately accounted for, limiting potential bias with respect to
	College or university education	33.2	35.3				
	European origin (%)	99.9	94.4				
	Married or cohabiting	89.5	91.4				
	Male sex baby	49.9	51.4				
	Inclusion criteria Births were identified using the Medical Birth Registry of Norway and the Norwegian Childhood Diabetes Registry. Gestational age was determined using the date of last menstrual period (LMP) or ultrasound-based estimations (where available), if LMP information was not available . If neither LMP nor ultrasound estimations were available, births were included in the study if the birthweight was greater than 500g. Exclusion criteria Births were excluded if there were no last menstrual period or						

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	ultrasound details (to establish gestational age) and the birthweight was less than 500g.				the prognostic factor of interest: Unclear 6) The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of invalid results: Yes
Full citation Kjos,S.L., Henry,O.A., Montoro,M., Buchanan,T.A., Mestman,J.H., Insulin-requiring diabetes in pregnancy: a randomized trial of active induction of labor and expectant management, American Journal	Sample size Participants were identified from the Women's Hospital, Los Angeles County-University of Southern California Medical Centre. Over 3000 women with diabetes were delivered during the study period of whom 944 required insulin therapy. 744 women did not meet the inclusion criteria or gestational diabetes was recently diagnosed or refused randomisation. n=200 Insulin dependent gestational	Interventions Active induction of labour at 38 weeks, n=100 Expectant management, n=100	Details Active induction of labour: In pregnancies where gestational age could not be determined with accuracy, amniocentesis was performed to assess foetal lung maturity. Women with 1) accurate estimation of gestational age or 2) evidence of foetal	Results Mode of delivery Caesarean section (operative indication - numbers in parentheses are those with caesarean section without labour) Elective induction group = 25/100 (Arrest disorder: 6, Failed induction of labour: 6, Foetal distress: 7 (2), Macrosomia: 1 (1), Elective repeat: 2 (2), Malpresentation: 3 (3)) Expectant management group = 31/100 (Arrest disorder: 12, Failed induction of labour: 8, Foetal distress: 3 (1), Macrosomia: 4 (3), Elective repeat: 3 (3), Malpresentation: 1) Caesarean section in women without previous	Limitations NICE guidelines manual. Appendix C: Methodology checklist: randomised controlled trials

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td>of Obstetrics and Gynecology, 169, 611-615, 1993
Ref Id
236279
Country/ies where the study was carried out
USA
Study type
Randomised controlled trial

Aim of the study
To assess if a program of expectant management of uncomplicated pregnancies of women with insulin-requiring gestational or pregestational class B diabetes would reduce caesarean birth incidence

Study dates
October 1987 to February 1991</td>
<td>diabetes = 187
Pregestational non-insulin dependent diabetes before pregnancy = 13
(9/13 in elective induction group, 4/13 in expectant management group)

Characteristics
100g OGTT used for diagnosis applying O'Sullivan or NDDG criteria

<table>
<tr><th></th><th>Active induction group Mean (95%CI)</th><th>Expectant management group Mean (95%CI)</th><th>p valu</th></tr>
<tr><td>Maternal age at delivery (yr)</td><td>32.1 (30.9-33.2)</td><td>31.9 (30.8-33.0)</td><td>NS</td></tr>
<tr><td>Gravidity</td><td>4.3 (3.9-4.7)</td><td>4.1 (3.7-4.5)</td><td>NS</td></tr>
<tr><td>Parity</td><td>2.5 (2.2-2.9)</td><td>2.4 (2.0-2.7)</td><td>NS</td></tr>
<tr><td>Maternal weight at delivery (kg)</td><td>83.7 (80.9-87.4)</td><td>85.0 (81.3-88.8)</td><td>NS</td></tr>
</table>
</td>
<td></td>
<td>lung maturity (lecithin sphingomyelin ratio ≥ 2.0) were scheduled within 5 days for induction of labour. If foetal lung maturity was not confirmed, amniocentesis was performed again 1 week later. Women continued twice weekly antepartum surveillance and hme insulin therapy. Labour was induced with intravenous oxytocin. Women with favourable Bishop scores (<4), unscarred uteri and normal amniotic fluid indices (>5.0cm), up to three applications of vaginal prostaglandin (3mg) were used for cervical ripening befor treatement with oxytocin.

Expectant management: Expectant management was</td>
<td>caesarean section
Elective induction group = 20/89 (22.5%)
Expectant management group = 14/80 (17.5%)
RR = 1.28 (95% CI 0.70 to 2.37)*

Vaginal delivery
Elective induction group = 75/100
Expectant management group = 69/100
RR = 1.09 (95%CI 0.91 to 1.29)*

Onset of labour
Spontaneous labour
Elective induction group = 22/100
Expectant management = 44/100

Induction of labour
Elective induction group= 70/100
Expectant management = 49/100
(indications for the 49 women were abnormal antenatal testing: 19, ruptured membranes without labour: 8, 42 gestational weeks: 7, poor foetal growth: 4, pregnancy induced hypertension: 3, suspected macrosomia: 1, maternal insistence on delivery:7)

Caesarean delivery without labour
Elective induction group= 8/100
Expectant management = 7/100 (One additional woman presented in spontaneous labour with a transverse foetal lie and underwent caesarean section without allowing labour to proceed)

Perinatal mortality (no congenital malformations in either group)
Elective induction group= 0/100</td>
<td>1) An appropriate method of randomisation was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups) - Unclear
2) There was adequate concealment of allocation (such that investigators, clinicians and participants cannot</td>
</tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Source of funding Not reported	<table><tr><td>Gestation at entry</td><td>38wk1d (38wk-38wk2d)</td><td>38wk2d (38wk1d-38wk3d)</td><td>NS</td></tr><tr><td>Interval to delivery (days)</td><td>6.4 (5.3-11.6)</td><td>12.8 (11.6-13.9)</td><td>0.0 1</td></tr><tr><td>Gestation at delivery (wks)</td><td>38wk (38wk6d - 39wk2d)</td><td>40wk (39wk6d - 40wk2d)</td><td>0.0</td></tr></table> Inclusion criteria Women diagnosed before pregnancy with insulin dependent diabetes mellitus or non-insulin dependent diabetes mellitus without vascular complications Women with gestational diabetes requiring insulin treatment during pregnancy and who had good metabolic control of blood glucose (assessed using capillary blood glucose self monitoring and defined as a preprandial or fasting blood glucose ≤90mg/dl and postprandial values ≤120mg/dl for 90% of readings) Further inclusion criteria for all women were 1) 38 gestational weeks completed 2) good compliance with clinic appointments and home glucose		daily split-dose insulin treatment and home blood glucose monitoring, weekly antenatal clinic appointments and twice weekly antepartum testing until spontaneous labour occurred. Induction of labour was undertaken if 1) decelerations or nonstress testing or low amniotic fluid volume indicated suspected foetal distress 2) preeclampsia occurred, 3) maternal hyperglycaemia or ketonuria occcured 4) estimated foetal weight ≥ 4200g or 5) the pregnancy exceeded 42 gestational weeks. Gestational age in both groups determined by last menstrual period adjusted if ultrasononographic estimation (before 22 weeks) indicated a	Expectant management = 0/100 RR = NC Neonatal hypoglycaemia requiring treatment (No definition given) Elective induction group= 0/100 Expectant management = 0/100 RR = NC Birth weight > 4000 g Elective induction group= 15/100 Expectant management = 27/100 RR = 0.56 (95%CI 0.32 to 0.98)* Birth weight > 4500 g Elective induction group= 0/100 Expectant management = 2/100 RR = 0.20 (95%CI 0.01 to 4.11)* Mild shoulder dystocia (no birth trauma - Erb's palsy or bone fracture - in either group) (No definition given) Elective induction group= 0/100 Expectant management = 3/100 RR = 0.14 (95%CI 0.01 to 2.73)*	influence enrolment or treatment allocation) - Unclear 3) The groups were comparable at baseline, including all major confounding and prognostic factors - Yes 4) The comparison groups received the same care apart from the intervention(s) studied - Yes 5) Participants receiving care were kept 'blind'

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	monitring 3) no abnormalities with non stress testing and amniotic fluid volume measurement performed from 34 gestational weeks onward as part of a twice weekly antenatal assessment 4) singleton gestation and cephalic presentation 5) clinical and ultrasonigraphic featal weight estimation ≤3800g at 38 completed gestational weeks with no evidence of intrauterine groth retardation 6) no other medical or obstetric complications 7) a candidate for trial of vaginal delivery (no more than 2 previous caesarean sections) Participants gave written informed consent. Exclusion criteria Not reported		difference of ≥ 10 days		to treatment allocation - No 6) Individuals administering care were kept 'blind' to treatment allocation - No 7) All groups were followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes 8) How many participants did not complete treatment

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					in each group? - None 9) The groups were comparable for treatment completion (that is, there were no important or systematic differences between groups in terms of those who did not complete treatment) - Yes 10) For how many participants in each group were no outcome data available? - None 11) The

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					groups were comparable with respect to the availability of outcome data (that is, there were no important or systematic differences between groups in terms of those for whom outcome data were not available) - Yes 12) The study had an appropriate length of follow-up - Yes 13) The study used a

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					precise definition of outcome - No definitions were given for shoulder dystocia or neonatal hypoglycaemia 14) A valid and reliable method was used to determine the outcome - Unclear for shoulder dystocia or neonatal hypoglycaemia 15) Investigators were kept 'blind' to

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					participants' exposure to the intervention - No 16) Investigators were kept 'blind' to other important confounding and prognostic factors - No Other information
Full citation Lurie,S., Insler,V., Hagay,Z.J., Induction of labor at 38 to 39 weeks of gestation reduces the incidence of shoulder dystocia in gestational diabetic patients class A2, American Journal of	Sample size 164 women with class A2 gestational diabetes met the criteria for enrollment in the period 1 January 1983 to 31 December 1989. 92 women with class A2 gestational diabetes met the criteria for enrollment in the period 1 January 1990 to 31 July 1994. Characteristics 1983 - 1989 | 1990	Interventions In the first period, unless foetal health was compromised, expectant mangement was observed.	Details In the first period, unless foetal health was compromised, pregnancy was allowed to progress to spontaneous labour. If the woman was undelivered at 40 gestational weeks a nonstress test and evaluation of cervical status were performed	Results Mode of delivery Caesarean section, n (%) Expectant management group = 31/164 (18.9%) Induction of labour group = 22/96 (22.9%) RR = 1.21 (95%CI 0.75 to 1.97)* Vacuum extraction, n (%) Expectant management group = 9/164 (5.5%) Induction of labour group = 5/96 (5.2%) RR = 0.95 (95%CI 0.33 to 2.75)* Spontaneous birth, n (%) Expectant management group = 128/164 (75.6%)	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies 1) Method

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
PerinatologyAm.J.Perinatol., 13, 293-296, 1996 Ref Id 240501 Country/ies where the study was carried out Israel Study type Prospective cohort study Aim of the study To examine whether shoulder dystocia would be significantly reduced by elective induction of labour at 38-39 gestational weeks in women with insulin requiring gestational diabetes (A2) Study dates Participants were recruited from two study periods - 1 January 1983 to 31 December 1989 and 1	<table><tr><td></td><td>protocol Expectant management</td><td>proto Indu labo</td></tr><tr><td>n</td><td>164</td><td>96</td></tr><tr><td>Mean maternal age (yr)</td><td>33.1 ± 5.0</td><td>32.5</td></tr><tr><td>Mean parity</td><td>2.5 ± 1.8</td><td>1.9 ±</td></tr><tr><td>Geational age at delivery (wk)</td><td>39.2 ± 1.6</td><td>38.4</td></tr><tr><td>Infant's weight at delivery (g)</td><td>3430.1 ±530.0</td><td>3406</td></tr></table> Inclusion criteria Women with gestational diabetes whose infants were delivered during both periods at the authors' maternal-foetal medical unit In the first period, gestational age was established on the basis of the last menstrual period and ultrasonographic crown-rump measurements in the first trimester. In the second period however, serial ultrasonographic crown-rump measurements were taken in the	In the second period, induction of labour was performed at 38 to 39 gestational weeks if appropriate.	twice weekly and biophysical score once a week. Induction of labour was attempted if one of the following was met. 1) Ultrasonographic estimation of an excessively large foetus (>4000g) 2) Assessment of biophysical score or OCT indicating compromise of foetal health 3) a Bishop score of >6 was obtained Instrumental delivery or caesarean section was perfomed as usually indicated. Elective caesarean section was performed where foetal weight was estimated to be ≥4500g. In the second period, an amniocentesis was performed to estimate lung maturity and the ratio of lectithin to sphingomyelin (L/S	Induction of labour group = 69/96 (71.9%) RR = 0.92 (95%CI 0.79 to 1.07)* Infants weighing >4000g, n (%) Expectant management group = 30/164 (18.3%) 15/30 delivered after 40 weeks Induction of labour group = 9/96 (9.4%) RR = 0.51 (95%CI 0.25 to 1.03)* Shoulder dystocia (corrected for caesarean delivery) (Definition: failure of the shoulder to be delivered spontaneously after the head due to impaction of the anterior shoulder against the symphasis pubis, as judged by the clinician delivering the foetus) Expectant management group = 7/133 (5.3%) 5/7 delivered after 40 weeks. 2/7 Erb's palsy, 1/7 clavicular fracture Induction of labour group = 1/74 (1.4%) this was a neonatal death due to asphyxia RR = 0.26 (95%CI 0.03 to 2.05)* Respiratory distress syndrome (No definition given) Expectant management group = 0/164 (0%) Induction of labour group = 0/96 (0%) RR = not calculable	of allocation to treatment groups was unrelated to potential confounding factors - Yes 2) Attempts were made within the design or analysis to balance the comparison groups for potential confounders - No 3) Groups were comparable at baseline, including all major confounding and

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
January 1990 to 31 July 1994 Source of funding None stated	first trimester. Exclusion criteria In both periods: 1) Multiple gestation pregnancy 2) Breech presentation 3) Complications of pre-eclampsia		ratio) and phosphatidylglycerol presence were assessed from the amniotic fluid. If the lungs were assessed to be mature and the cervix was unfavourable (Bishop score <6), induction of labour was performed by either intracervical balloon catheter or placement of 0.5mg prostaglandin E2 gel. If the cervix was favourable, intravenous oxytocin was administered followed by amniotomy. If foetal weight was estimated to be ≥4500g by clinical or ultrasound examination, the mother was delivered by caesarean section.		prognostic factors - Yes 4) Comparison groups received the same care apart from the intervention(s) studied - Yes 5) Participants receiving care were kept 'blind' to treatment allocation - No 6) Individuals administering care were kept 'blind' to treatment allocation - No 7) All groups were

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					followed up for an equal length of time (or analysis was adjusted to allow for differences in length of follow-up) - Yes 8) How many participants did not complete treatment in each group? - None 9) Groups were comparable for treatment completion - Yes 10) For how many participants in each group were no

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					outcome data available? - None 11) Groups were comparable with respect to the availability of outcome data - Yes 12) The study had an appropriate length of follow-up - Yes 13) The study used a precise definition of outcome - Yes for shoulder dystocia but not for respiratory distress 14) A valid

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					and reliable method was used to determine the outcome - Yes 15) Investigators were kept 'blind' to participants' exposure to the intervention - No 16) Investigators were kept 'blind' to other important confounding and prognostic factors - No Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation Alberico,S., Businelli,C., Wiesenfeld,U., Erenbourg,A., Maso,G., Piccoli,M., Ronfani,L., Gestational diabetes and fetal growth acceleration: induction of labour versus expectant management, Minerva Ginecologica, 62, 533-539, 2010 Ref Id 236644 Country/ies where the study was carried out Trieste, Italy Study type Retrospective cohort study Aim of the study To compare elective induction of labour at 38 gestational weeks	Sample size 230 women diagnosed with gestational diabetes at the Maternal and Child Institute IRCCS Burlo Garofalo between 1996 and 2007 of whom 99 were eligible for inclusion to the study Characteristics [see table below]	Interventions Intervention: elective induction of labour was performed by administration of PGE2 gel every 6-8 hours until labour started. If induction did not succeed after 5 attempts then caesarean section was performed. Control: women in the expectant management group were reassesse	Details 99 women were included in the study. 48 women underwent induction of labour and 51 were managed expectantly.The primary outcome was caesarean section rate and secondary outcomes were macrosomia, neonatal Apgar score, NICU admissions, shoulder dystocia and perinatal mortality.	Results Mode of delivery Caesarean section Elective induction group: 9/48 (19%), 8/9 failed induction, 1/9 foetal distress Expectant management group: 11/51 (22%), 8/11 macrosomia, 2/11 foetal distress, 1/11 following induction>38 weeks RR = 0.87 (95% CI 0.40 to 1.91)* Subgroup of women with normal BMI (20-25) Elective induction group: 14% Expectant management group: 14% OR = 0.99 (95% CI 0.2 to 4.91) Subgroup of women with obesity (BMI ≥30) Elective induction group: 24% Expectant management group: 50% OR = 0.31 (95%CI 0.04 - 2.14) Comparison of obese vs normal weight women Obese women = 33% Normal weight women = 14% p=0.03 Multivariate analysis of women with BMI ≥30 vs women with BMI <30 Adjusted OR = 3.9 (95% CI 1.2 to 12.8) (adjusted for maternal age, parity, hypertensive disorders and induction of labour at 38 gestational weeks) Operative delivery Elective induction group: 3/48 (6%) Expectant management group: 1/51 (2%) RR = 3.19 (95% CI 0.34 to 29.60) Spontaneous delivery	Limitations NICE guidelines manual. Appendix D: Methodology checklist: Cohort studies 1) Method of allocation to treatment groups was unrelated to potential confounding factors - Unclear 2) Attempts were made within the design or analysis to balance the compariso

	Induction at 38 weeks (N=48)	Exp ma (N=
Age (years) Mean ± SD	33.3 ± 4.9	32.
Nulliparas (%)	30 (63%)	30
Mean maternal BMI	28 ± 7	25
<25	37%	50%
25-29	26%	28%
30-34	20%	17%
≥35	17%	4%
Obesity (BMI≥30)	37%	21%
Positive urine protein test	28 (58%)	24
Insulin therapy	8 (17%)	5 (1
Ketonuria	9 (19%)	7 (1

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
with expectant management in women with gestational diabetes (A1 and A2) and foetal growth acceleration Study dates Between 1996 and 2007 Source of funding None stated	<table><tr><td>Hypertension (≥140/90mmHg)</td><td>10 (21%)</td><td>10</td></tr><tr><td>Impaired glycaemic profile</td><td>17 (42%) available for 41/48</td><td>8(2 ava for</td></tr></table> Inclusion criteria Women with gestational diabetes and foetal growth acceleration diagnosed at 38 gestational weeks Diagnosis of gestational diabetes was based on: 1) a positive 50g glucose challenge test (≥140mg/dl) between 24 and 28 weeks 2) if the 50g glucose challenge test result was 140 - 184 mg/dl, then a 75g OGTT was performed. If at least 2/3 results were above threshold (Fasting 95mg/dl, 1 hr 180 mg/dl, 2 hr 155mg/dl) then a positive diagnosis was made 3) if the 50g glucose challenge test result was ≥185mg/dl, then a diagnosis of gestational diabetes was made without further testing Foetal monitoring was by monthly ultrasound assessment from 28-30 gestational weeks. Acceleration was defined by a foetal growth	d at 40-41 gestational weeks by ultrasound. If the estimated foetal weight was >4250g, then a caesarean section was performed, otherwise the patient was observed until spontaneous labour started. Induction was offered if there were any new emerging indications (oligohydramnios, PROM, post-term pregnancy		Elective induction group: 36/48 (75%) Expectant management group: 39/51 (76%), 3/39 following induction>38 weeks RR = 0.98 (95% CI 0.78 to 1.23)* Induction > 38 weeks in expectant management group: 4/51 (8%) for reasons not related to gestational diabetes, 3/4 spontaneous delivery, 1/4 caesarean section Macrosomia (Definition: Birthweight >4000g) Elective induction group: 6/48 (13%) Expectant management group: 11/51 (22%) p=0.2 RR = 0.58 (95% CI 0.23 to 1.44)* Admission to NICU (No definition given) Elective induction group: 1/48 (2%) Expectant management group: 6/51 (12%) p=0.1 RR = 0.18 (95% CI 0.02 to 1.42)* Shoulder dystocia (No definition given) Elective induction group: 0/48 (0%) Expectant management group: 0/51 (0%) RR = NC Stillbirth (No definition given) Elective induction group: 0/48 (0%) Expectant management group: 1/51 (2%) RR = 0.35 (95% CI 0.01 to 8.48)	n groups for potential confounders - Yes 3) Groups were comparable at baseline, including all major confounding and prognostic factors - Yes although there were significantly more very obese women in the elective delivery group compared to the expectant management group 4) Comparison groups

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	exceeding 2SDs of the expected values of common ultrasound measurements (crown-rump length, head circumference, abdominal circumference and femoral length) at 38 gestational weeks Exclusion criteria 1) An estimated foetal weight ≥4250g 2) Presence of another indication for elective caesarean section 3) Previous caearean delivery	). For both groups, a caesarean section was performed if foetal distress was suspected.			received the same care apart from the intervention(s) studied - Yes 5) Participants receiving care were kept 'blind' to treatment allocation - No 6) Individuals administering care were kept 'blind' to treatment allocation - No 7) All groups were followed up for an equal length of time (or analysis

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					was adjusted to allow for differences in length of follow-up) - Yes 8) How many participants did not complete treatment in each group? - None 9) Groups were comparable for treatment completion - Yes 10) For how many participants in each group were no outcome data available? - None 11) Groups

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					were comparable with respect to the availability of outcome data - Yes 12) The study had an appropriate length of follow-up - Yes 13) The study used a precise definition of outcome - Unclear for most outcomes, definition only given for macrosomia 14) A valid and reliable method was used

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					to determine the outcome - Unclear for still birth and admission to NICU 15) Investigators were kept 'blind' to participants' exposure to the intervention - No 16) Investigators were kept 'blind' to other important confounding and prognostic factors - No Other information

A.17 Diagnostic accuracy and timing of postnatal testing

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation Vambergue,A., Dognin,C., Boulogne,A., Rejou,M.C., Biausque,S., Fontaine,P., Increasing incidence of abnormal glucose tolerance in women with prior abnormal glucose tolerance during pregnancy: DIAGEST 2 study, Diabetic Medicine, 25, 58-64, 2008 Ref Id 116599 Country/ies where the study was carried out France Study type Prospective cohort	Sample size Number with gestational diabetes: 466 Number with postnatal test: FPG (295/466, 63.3%) OGTT (209/466, 44.8%) Characteristics Maternal age in years, mean (SD) In subjects with normal glucose tolerance at follow-up: 37.0 (5.6) In subjects with IFG at follow-up: 38.8 (6.7) In subjects with IGT at follow-up: 39.2 (5.8) In subjects with diabetes at follow-	Tests 75g 2 hour OGTT	Methods -Gestational diabetes criteria: 50g glucose challenge test. If the 1 hour value was ≥ 7.2mmol/l, then a 100g 3 hour OGTT was performed. Women who had 2 or more of the four OGTT values above Carpenter and Coustan's criteria (fasting ≥5.3mmol/l, 1 hour ≥10.0mmol/l, 2 hour ≥8.6mmol/l and 3 hour ≥7.8mmol/l) were defined as having gestational diabetes -Outcomes: Diabetes, IFG, IGT -Outcome definitions: ADA criteria. Diabetes was defined as FPG ≥7.0mmol/l or a 2 hour glucose ≥11.1mmol/l. IGT was defined as FPG <7.0mmol/l and 2 hour ≥7.8 but	Results Incidence data IGT: 13.4% (28/209) - based on OGTT measurements Diabetes: 18% (53/295) - based on FPG measurements	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: No 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
study Aim of the study To determine the prevalence of diabetes, impaired glucose tolerance or impaired fasting glucose 6.75 years after delivery in women with differential blood glucose status during pregnancy Study dates NR Source of funding Research was supported by the pharmaceutical firms Lifescan and NovoNordisk	up: 39.6 (6.4) Ethnicity, % French In subjects with normal glucose tolerance at follow-up: 95.4 In subjects with IFG at follow-up: 85.7 In subjects with IGT at follow-up: 72.1 In subjects with diabetes at follow-up: 75.8 Parity, mean (SD) NR Family history of diabetes, % In subjects with normal glucose tolerance at follow-up: 76.1 In subjects with IFG at follow-up: 72.2 In subjects with IGT at follow-up: 71.8 In subjects with diabetes at follow-up: 54		<11.1mmol/l. IFG was defined by FPG ≥5.6mmol/l but <7.0mmol/l. -Timing of postnatal test: 6 years -Location of postnatal test (primary/secondary care): Laboratory -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information Only data for diabetes and IGT have been extracted as cut-off for IFG does not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	BMI, kg/m2, mean (SD) In subjects with normal glucose tolerance at follow-up: 27.1 (5.7) In subjects with IFG at follow-up: 32.6 (8.0) In subjects with IGT at follow-up: 30.5 (7.4) In subjects with diabetes at follow-up: 32.3 (6.8) Macrosomia (%) NR Medication during pregnancy, % insulin NR * The characteristics above are of those who completed the postnatal test Inclusion Criteria Women with				

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	gestational diabetes recruited from 15 public maternity units in northern France Exclusion Criteria NR				
Full citation Albareda,M., Caballero,A., Badell,G., Piquer,S., Ortiz,A., de,Leiva A., Corcoy,R., Diabetes and abnormal glucose tolerance in women with previous gestational diabetes, Diabetes Care, 26, 1199-1205, 2003 Ref Id 152953 Country/ies where the study was carried out Spain Study type Retrospective cohort study	Sample size Number with gestational diabetes: 982 Number with postnatal test: 696 Characteristics Maternal age in years, median (range) 31(17-44) Ethnicity Spanish women Parity, mean (SD) Not reported Family history of diabetes, % 373/695 (53.7%)	Tests 2 hour 75g OGTT	Methods -Gestational diabetes criteria: 50g 1 hour glucose challenge test. Criteria for screening and glucose tolerance testing were those from the Second and Third Workshop Conferences on gestational diabetes -Outcomes: Diabetes, IFG, IGT -Outcome definitions: WHO 1999 (cut-offs not reported in article) -Timing of postnatal test: 6 weeks after delivery or after cessation of breast feeding, whichever occurred later. A second test 5 years after the first -Location of postnatal test	Results Incidence data At 6 years Diabetes: 5.6% (39/696) IGT: 8.8% (61/696) IFG: 3.6% (25/696) At 11 years Diabetes: 13.8% (NR/NR)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Aim of the study To assess the progression to diabetes and abnormal glucose tolerance (AGT) of Spanish women with gestational diabetes and to identify predictive factors Study dates All women were diagnosed with gestational diabetes between 1986 and 1993 Source of funding Not reported	Prepregnancy BMI, kg/m2, median (range) 23.3 (15.9-37.9) Macrosomia (%) 25/692 (3.6%) Medication during pregnancy, % insulin 472/695 (67.9%) * The characteristics above are of those who completed the postnatal test Inclusion Criteria Women with gestational diabetes who attended the Diabetes and Pregnancy Clinic Exclusion Criteria NR		(primary/secondary care): Unclear -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported
Full citation Buchanan,T.A., Xiang,A.,	Sample size Number with gestational	Tests 75g OGTT	Methods -Gestational diabetes criteria: Recommendations o	Results Incidence data	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Kjos,S.L., Lee,W.P., Trigo,E., Nader,I., Bergner,E.A., Palmer,J.P., Peters,R.K., Gestational diabetes: antepartum characteristics that predict postpartum glucose intolerance and type 2 diabetes in Latino women, Diabetes, 47, 1302-1310, 1998 Ref Id 153030 Country/ies where the study was carried out USA Study type Prospective cohort study Aim of the study To examine antenatal clinical characteristics along with measures of glucose tolerance, insulin sensitivity,	diabetes: 233 Number with postnatal test: 122 (52%) Characteristics Maternal age in years (antenatally), mean (SD) Of those with normal glucose tolerance postpartum: 30.8 (5.2) Of those with IGT postpartum: 29.3 (5.7) Of those with diabetes postpartum: 32.3 (6.2) Ethnicity All Latino women Parity, mean (SD) NR Family history of diabetes, % NR BMI, kg/m2, mean (SD)		f the Third International Workshop Conference on Gestational diabetes; measurement of the plasma glucose concentration 1 hour after ingestion of 50g glucose. Women with a value ≥7.8mmol/l underwent a 3 hour 100g OGTT to make or exclude the diagnosis of gestational diabetes -Outcomes: IGT, diabetes -Outcome definitions: ADA 1997 criteria. Cut-offs not reported in article but extracted from a reference article. IGT defined as 2 hour glucose ≥7.8 and <11.1mmol/l and diabetes defined as fasting ≥7 or 2 hour ≥11.1mmol/l -Timing of postnatal test: 1-6 months -Location of postnatal test (primary/secondary care): Unclear -Did study document a return to euglycaemia in the immediate days following delivery and before	Diabetes: 12/122 (10%)	diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
pancreatic B-cell function, and body composition in Latino women with gestational diabetes for their ability to predict type 2 diabetes or impaired glucose tolerance within 6 months of delivery Study dates August 1993-March 1995 Source of funding Grants from the National Institute of Diabetes, Digestive and Kidney Diseases of the National Institutes of Health; grants from the General Clinical Research Center Branch of the National Institutes of Health; and the Medical Research Service of the Department of Veterans Affairs	Prepregnancy BMI Of those with normal glucose tolerance postpartum: 30.4 (5) Of those with IGT postpartum: 28.0 (4.1) Of those with diabetes postpartum: 29.1 (4) Postpartum BMI Of those with normal glucose tolerance postpartum: 30.8 (4.9) Of those with IGT postpartum: 29.4 (4.8) Of those with diabetes postpartum: 29.5 (3.5) Macrosomia (%) NR Medication during pregnancy, % insulin Of those with normal glucose		discharge: No		results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information Only data for diabetes has been extracted as cut-off for IGT does not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	tolerance postpartum: 8.2 Of those with IGT postpartum: 18.7 Of those with diabetes postpartum: 8.3 * The characteristics above are of those who completed the postnatal test Inclusion Criteria - Between 29 and 34 weeks gestation as assessed by a clinical examination before 12 weeks' gestation or an ultrasound before 20 weeks' gestation - Women not on insulin therapy - Women with fasting serum glucose concentrations				

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	<7.2mmol/l since the diagnosis of gestational diabetes - Women with otherwise uncomplicated singleton pregnancies - Only women whose parents and at least three of four grandparents were from Mexico, Guatemala or El Salvador were recruited Exclusion Criteria - 3 of the 153 women who came for antenatal testing had circulating anti-islet cell antibodies and were excluded				
Full citation Jang,H.C., Yim,C.H.,	Sample size Number with gestational	Tests 2-hour 75g	Methods -A prospective study which performed 75g OGTTs	Results Incidence data	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Han,K.O., Yoon,H.K., Han,I.K., Kim,M.Y., Yang,J.H., Cho,N.H., Gestational diabetes mellitus in Korea: prevalence and prediction of glucose intolerance at early postpartum, Diabetes Research and Clinical Practice, 61, 117-124, 2003 Ref Id 153332 Country/ies where the study was carried out Korea Study type Prospective cohort study Aim of the study To determine the prevalence of glucose intolerance in Korean women with gestational diabetes between	diabetes: 392 Number with postnatal test: 311 (79%) Characteristics Maternal age in years, mean (SD) 30.9 (4.1) Race/ethnicity Korean women Parity, mean (SD) 0.5 (0.7) Family history of diabetes (%) 40.5 Prepregnancy BMI, kg/m2: 22.7 (3.5) Macrosomic infant delivered NR Insulin use during pregnancy (%)	OGTT	between 6 and 8 weeks' postpartum in women with gestational diabetes -Gestational diabetes criteria: Women with a positive screen (plasma glucose concentrations >=7.2mmol/l, 1 hour after 50g glucose load) were recalled for a 3-hour, 100g OGTT within 2 weeks. Women were considered to have gestational diabetes if at least two values reached or exceeded the following thresholds: 5.8mmol/l at fasting, 10.6mmol/l at 1 hour, 9.2mmol/l at 2 hours, 8.1mmol/l at 3 hours; NDDG criteria -Outcomes: Diabetes, IGT -Outcome definitions: ADA 1997. Cut-offs not reported in article but extracted from reference given for diabetes: FPG >=7mmol/l or 2-hour PG >=11.1mmol/l. IGT: 2-hour PG>=7.8 and <11.1mmol/l) -Timing of postnatal test: 6-8 weeks after delivery	Diabetes: 47/311 (15.1%)	diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
6 and 8 weeks postpartum and identify which antenatal clinical and metabolic variables were predictive of postpartum diabetes and impaired glucose tolerance (IGT) Study dates All women were screened for gestational diabetes between January 1993 and June 1997 Source of funding NR	NR * The characteristics above are of those who completed the postnatal test Inclusion Criteria -Women with gestational diabetes and follow-up evaluation of glucose intolerance between 6 and 8 weeks postpartum Exclusion Criteria -Subsequent pregnancies in women with gestational diabetes		-Location of postnatal test (primary/secondary care): unclear -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information -NR: Not reported -Only data for diabetes was extracted as cut-off for IGT in this article does not match the WHO criteria
Full citation Kwong,S., Mitchell,R.S., Senior,P.A., Chik,C.L., Postpartum diabetes	Sample size Number with gestational diabetes: 909 Number with postnatal test:	Tests 75g 2-hour OGTT FPG only: 21/438	Methods - Retrospective cohort study of women with gestational diabetes attending a pregnancy diabetes clinic. Data were obtained from patient medical records	Results Incidence data Type 2 diabetes: 14/438 (3%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
screening: adherence rate and the performance of fasting plasma glucose versus oral glucose tolerance test, Diabetes Care, 32, 2242-2244, 2009 Ref Id 153432 Country/ies where the study was carried out Canada Study type Retrospective cohort study Aim of the study To determine the rate of adherence to postnatal glycaemic testing in women with gestational diabetes and the performance of FPG versus the 75g OGTT in detecting postnatal glucose intolerance	438 (48.2%) Characteristics Age in years, mean (SD) 32.0 +/-4.5 Ethnicity, n (%) Caucasian: 247 (56.4) Non-caucasian: 190 (43.4) Parity, mean (SD) 0.87 +/-0.97 Family history of diabetes, n (%) Present: 286 (65.3) Absent: 147 (33.6) Prepregnancy BMI (kg/m2), mean (SD) 27.7 +/-6.2 Macrosomic infant delivered NR	(5%) OGTT: 417/438 (95%)	- Gestational diabetes criteria: A 1-hour plasma glucose measurement after a 50g glucose load of >=10.3mmol/l was considered as diagnostic of gestational diabetes, and <7.8mmol/l was considered normal. 75g OGTT was undertaken in women in between these two values. Two or more abnormal values (FPG >=5.3mmol/l, 1-hour plasma glucose >=10.6mmol/l and 2-hour plasma glucose >=8.9mmol/l) diagnostic of gestational diabetes - Canadian Diabetes Association (CDA) criteria -Outcomes: Type 2 diabetes, IFG, IGT. -Outcome definitions: diabetes was defined as FPG >=7mmol/l or 2-hour plasma glucose >=11.1mmol/l, IFG as FPG of 6.1-6.9mmol/l and IGT as 2-hour plasma glucose of 7.8-11.1mmol/l (CDA criteria). -Timing of postnatal test: 6		2) Were selection criteria clearly described: No 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Study dates Women seen at clinic between April 1999 and March 2006 Source of funding NR	Insulin use during pregnancy, n (%) Present: 287 (65.6%) Absent: 146 (33.3%) * The characteristics above are of those who completed the postnatal test Inclusion Criteria -All consecutive women with gestational diabetes or IGT of pregnancy Exclusion Criteria -Women with pre-existing hyperglycaemia (type 1 or type 2 diabetes, IFG or IGT) and those who did not undergo routine screening for gestational diabetes		weeks - 6 months -Location of postnatal test (primary/secondary care): NR -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information -NR: Not reported -Data for only diabetes have been extracted as the cut-off for other outcomes in this article do not match the WHO 1999 criteria.

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation Lauenborg,J., Hansen,T., Jensen,D.M., Vestergaard,H., Molsted-Pedersen,L., Hornnes,P., Locht,H., Pedersen,O., Damm,P., Increasing incidence of diabetes after gestational diabetes: a long-term follow-up in a Danish population, Diabetes Care, 27, 1194-1199, 2004 Ref Id 153456 Country/ies where the study was carried out Denmark Study type Retrospective cohort study Aim of the study To study the incidence of	Sample size Number with gestational diabetes: 753 (241 from old cohort, 512 from new cohort) Number with postnatal test: 481/753 (63.9%) Characteristics Age at index pregnancy in years, median (IQR) 31.7 (27.7-35.7) Ethnicity Danish population Parity, mean (SD) Not reported Family history of diabetes, n(%) Not reported Prepregnancy BMI (kg/m2), median (IQR) 25.1 (21.9-29.8) Macrosomic	Tests 2 hour 75g OGTT (5% of tests were based on capillary whole blood glucose due to technical problems obtaining venous samples)	Methods -Women with diet-treated gestational diabetes during 1978-1985 (old cohort, n=241, also followed up around 1990) or 1987-1996 (new cohort, n=512) were examined in 2000-2002. Women were classified by a 2 hour 75g OGTT according to the WHO criteria or an intravenous glucagon test supplemented by measurement of Glutamic Acid Decarboxylase (GAD) antibodies. Historical data from index-pregnancy and anthropometrical measurements were collected. 64% (n=481; 151/241 of old cohort, 330/512 of new cohort) of the total population was included -Gestational diabetes criteria: OGTTs were defined as abnormal if two or more of 7 values during the test exceeded 3 SDs above the mean for a group of normal weight nonpregnant women without family history of diabetes examined in exactly	Results Incidence data Diabetes: 171/481 (36%) IGT/IFG: 130/481 (27%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
diabetes among women with previous diet-treated gestational diabetes in the light of the general increasing incidence of overweight and diabetes and to identify risk factors for the development of diabetes Study dates Women with gestational diabetes during 1978-1985 (old cohort) or 1987-1996 (new cohort) were examined in 2000-2002 Source of funding This research was supported by the Danish Medical Research Council, Copenhagen University, the Danish Diabetes Association, Handelsgartner	infant delivered Not reported Insulin use during pregnancy, n(%) Not reported * The characteristics above are of those who completed the postnatal test Inclusion Criteria - "Old cohort" comprised 241 women from the center for diabetes and pregnancy, Rigshopitalet, with diet-treated gestational diabetes during 1978-1985 who previously participated in a follow-up 2-11 years after index pregnancy. All subjects had gestational diabetes based		the same manner-fasting venous plasma glucose 6.4 and 6.2 and 2 hour plasma glucose 7.6 and 8.9mmol/l, respectively -Outcomes: diabetes, IFG/IGT -Outcome definitions: WHO 1999 criteria. Cut-off levels not reported in article but extracted from report of WHO/IDF consultation. IFG: FPG >=6.1 and <7mmol/l and 2 hour glucose <7.8mmol/l if measured. IGT: FPG<7.0mmol/l and 2 hour PG >=7.8 and <11.1mmol/l. Diabetes: FPG>=7.0 or 2 hour PG >=11.1mmol/l. -Timing of postnatal test: 2 months postpartum and subsequently in 1 to 2 year intervals, unless diabetes was diagnosed -Location of postnatal test (primary/secondary care): Center for diabetes and pregnancy, Rigshospitalet -Did study document a return to euglycaemia in the		10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: Yes Other information

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Ove Viliam Buhl Olesen og aegtefaelle Edith Buhl Olesens Mindelegat and Dagmar Marshalls Fond	on a 3 hour, 50g OGTT during pregnancy. -"New cohort" comprised all women (n=512) from the same center with diet-treated gestational diabetes between 1987 and 1996. Gestational diabetes diagnosis was based on a 3 hour, 75g OGTT. Exclusion Criteria -During 1986, the 50g OGTT was replaced by a 75g test, and women from 1986 were not included in the present follow-up study		immediate days following delivery and before discharge: No		
Full citation Lee,H., Jang,H.C., Park,H.K., Metzger,B.E., Cho,N.H.,	Sample size -Number with gestational diabetes: 868	Tests 75g 2-hour OGTT	Methods -The analysis included 620 gestational diabetes subjects. The postnatal examination	Results Incidence data, n(%) based on FPG alone Diabetes in the cases	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Prevalence of type 2 diabetes among women with a previous history of gestational diabetes mellitus, Diabetes Research and Clinical Practice, 81, 124-129, 2008 Ref Id 153463 Country/ies where the study was carried out Korea Study type Case-control study Aim of the study To determine whether Korean women with a history of gestational diabetes are at greater risk of developing type 2 diabetes than the general population Study dates Subjects recruited between August	-Number with postnatal test: 620 (71.4%) Characteristics Age in years, mean (SD) 33.6 (4.7) Ethnicity, n(%) NR Parity NR Family history of diabetes (% yes) 36.5 BMI, mean (SD) 23.5 (3.5) Macrosomic infant delivered NR Medication use during pregnancy NR		included a 2-hour 75g OGTT, lipid profiles, anthropometric measurements, and documentation of medical history, diet and lifestyle. All participants were followed up at 6 weeks postpartum and then annually. General population subjects were identified from the 2001 Korean National Health and Nutrition Survey and age-matched for case-control analysis -Gestational diabetes criteria: NDDG criteria- A 50g glucose challenge test was performed during 24-28 weeks' gestation. If the 1-hour plasma glucose value was >= 130mg/dl (7.2mmol/l), a 3-hour OGTT was conducted at 28-32 weeks' gestation. Cut-offs for gestational diabetes not reported in article but extracted from reference given: ≥2 glucose values (venous plasma) at or exceeding the following thresholds after a 100g OGTT: fasting, 105 mg/dl (5.8mmol/l); 1 hour, 190 mg/dl (10.6mmol/l); 2 hours,	(gestational diabetes subjects): 71/620 (11.5%)	1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (inclusion and exclusion criteria not reported) 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (though whole sample had OGTT, only FPG value was used to define diabetes) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
1995 and May 1997 Source of funding Supported by a Korean Science and Engineering Foundation Special Basic Research Grant	Inclusion Criteria Women with history of gestational diabetes Exclusion Criteria - Subjects with missing data - Subjects from the gestational diabetes 'B1' group (fasting glucose >=7.2mmol/l before the OGTT at 28-32 weeks of gestation) who may have had undiagnosed diabetes before pregnancy		165 mg/dl (9.2mmol/l); and 3 hours, 145 mg/dl(8.1mmol/l) -Outcomes: Diabetes -Outcome definitions: diabetes was diagnosed by a fasting plasma glucose >=7mmol/l (126mg/dl)*. Though gestational diabetes subjects underwent a 2-hour 75g OGTT during subsequent follow-ups, only the fasting plasma glucose value was used to define diabetes *Article does not state which criteria this is. Cut-off matches WHO 1999, ADA 1997, ADA 2003 and CDA -Timing of postnatal test: 6 weeks -Location of postnatal test (primary/secondary care): secondary care (study was conducted at 3 university hospitals-assuming women returned for follow-up postnatal test at same location) -Did study document a return		11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Unclear (not all clinical characteristics were reported) 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA (no withdrawals) Other information Data only extracted for the cases (women with gestational diabetes) NR: Not reported

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
			to euglycaemia in the immediate days following delivery and before discharge: No		
Full citation Lin,C.H., Wen,S.F., Wu,Y.H., Huang,Y.Y., Huang,M.J., The postpartum metabolic outcome of women with previous gestational diabetes mellitus, Chang Gung Medical Journal, 28, 794-800, 2005 Ref Id 153478 Country/ies where the study was carried out Taiwan Study type Prospective cohort study Aim of the study To determine the postnatal metabolic	Sample size Number with gestational diabetes: 235 Number with postnatal test: 127 (54%) Characteristics Age in years, mean (SD) 33.7 (4.1) Ethnicity Not reported Parity, mean (SD) 1.7 (0.9) Family history of diabetes, % 69.3 Prepregnancy BMI (kg/m2), mean (SD) 22.4 (3.7)	Tests 75g OGTT	Methods -From March 2001 to February 2003, 127 prior gestational diabetes women underwent a 75g OGTT and metabolic assessment at least 6 weeks after delivery. To identify the predictors, clinical variables obtained at the time of gestational diabetes were compared -Gestational diabetes criteria: Subjects were screened at 24-28 weeks' gestation and diagnosis of gestational diabetes was based on a 50g glucose challenge test of 1-hour plasma glucose level >=140mg/dl, followed by at least two abnormal values in a 100g OGTT. Women with documented gestational diabetes fulfilled the Carpenter and Coustan modification of the NDDG criteria (requiring at least two of the following: fasting glucose >=95, 1 hour>=180,	Results Incidence data Diabetes: 17/127 (13.4%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
abnormalities and predictive factors for subsequent diabetes in prior-gestational diabetes women in Taiwan Study dates All women with gestational diabetes diagnose d from March 2001 to February 2003 Source of funding This work was supported by a grant from Chang Gung Memorial Hospital	Prior macrosomia, % Not reported Medication use,% insulin Not reported * The characteristics above are of those who completed the postnatal test Inclusion Criteria -Women diagnosed with gestational diabetes at Tapei Chang-Gung Memorial Hospital. No women had a history of diabetes before pregnancy Exclusion Criteria -Not reported		2 hour>=155, 3 hour >=140mg/dl) -Outcomes: normal glucose tolerance, abnormal glucose tolerance (IFG or IGT), diabetes. ADA 1997 criteria-cut offs not reported but extracted from a reference article -Outcome definitions: normal was defined as fasting <6.1mmol/l and 2 hour <7.8mmol/l, IFG was defined as fasting ≥6.1mmol/l and <7.0mmol/l, IGT was defined as 2 hour ≥ 7.8 and <11.1mmol/l, diabetes was defined as fasting ≥7mmol/l or 2 hour ≥11.1mmol/l. -Timing of postnatal test: 1-19 months after delivery -Location of postnatal test (primary/secondary care): Taipei Chang-Gung Memorial Hospital -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported Only data for diabetes has been extracted as this matches WHO.

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation Lobner,K., Knopff,A., Baumgarten,A., Mollenhauer,U., Marienfeld,S., Garrido-Franco,M., Bonifacio,E., Ziegler,A.G., Predictors of postpartum diabetes in women with gestational diabetes mellitus, Diabetes, 55, 792-797, 2006 Ref Id 153484 Country/ies where the study was carried out Germany Study type Prospective cohort study Aim of the study To stratify risk for postnatal diabetes in women who have had gestational diabetes	Sample size Number with gestational diabetes: NR Number with postnatal test: 302 participated in follow-up, cumulative drop-out rate was 21% by 5 years Characteristics Maternal age at delivery in years, median (IQR) In islet autoantibody-positive women: 29.9 (27.5-31.7) In islet autoantibody-negative women: 31.4 (28.2-32.8) Ethnicity NR Parity, n (%) In islet autoantibody-negative women None: 125/270 (46)	Tests 75g 2 hour OGTT	Methods -Gestational diabetes criteria: German Diabetes Association using an OGTT with 75g glucose. Gestational diabetes was diagnosed if two of three capillary blood glucose values exceeded the following limits: >5mmol/l (fasting) before OGTT, >10.6mmol/l after 60 minutes, and >8.9mmol/l after 120 minutes. -Outcomes: Diabetes -Outcome definitions: ADA criteria. Cut-offs not reported in article but extracted from a reference article. Diabetes defined by FPG ≥7.0mmol/l or 2 hour glucose ≥11.1mmol/l -Timing of postnatal test: 9 months, 2, 5, 8 and 11 years -Location of postnatal test (primary/secondary care): Not reported -Did study document a return to euglycaemia in the immediate days following delivery and before	Results Incidence data 8 year cumulative risk of diabetes: 52.7% (55*/105) *Numerator not reported but estimated by NCC-WCH technical team	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Study dates Between 1989 and 1999, women with gestational diabetes were recruited from hospitals in Germany Source of funding Supported by grants from the German Federal Ministery for Education and Research and the German Diabetes Association and by a Federation of European Biochemical Societies fellowship to one of the authors	1: 88/270 (33) 2: 36/270 (13) >2: 21/270 (8) Data for islet autoantibody-positive women not reported Family history of diabetes, n(%) no/yes In islet autoantibody-negative women No: 155/253 (61) Yes: 98/253 (39) BMI, kg/m2, median (IQR) In islet autoantibody-positive women: 22.9 (21.1-25.7) In islet autoantibody-negative women: 26.5 (23.0-30.8) Macrosomia (%) NR Medication during pregnancy, n (%) insulin		discharge: Yes		without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	In islet autoantibody-positive women: 24/32 (75) In islet autoantibody-negative women: 92/270 (34.1) Inclusion Criteria - Women with gestational diabetes recruited from hospitals in Germany Exclusion Criteria NR				
Full citation Noussitou,P., Monbaron,D., Vial,Y., Gaillard,R.C., Ruiz,J., Gestational diabetes mellitus and the risk of metabolic syndrome: a population-based study in Lausanne, Switzerland,	Sample size Number with gestational diabetes: 159 Number with postnatal test: 74 (46.5%) Characteristics Maternal age in years at diagnosis of gestational diabetes, mean (SD)	Tests 2 hour 75g OGTT	Methods -Gestational diabetes criteria: Women with one or more risk factors for gestational diabetes underwent a 100g 3 hour OGTT. The diagnosis of gestational diabetes was made according to the NDDG criteria (≥2 abnormal values): ≥5.8mmol/l for fasting, ≥10.6mmol/l at 1 hour, ≥9.2mmol/l at 2 hours and ≥8.1mmol/l at 3 hours	Results Incidence data IGT: 16% (12/74) Diabetes: 11% (8/74)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Diabetes and Metabolism, 31, 361-369, 2005 Ref Id 153585 Country/ies where the study was carried out Switzerland Study type Retrospective cohort study Aim of the study To investigate the relationships between gestational diabetes and the metabolic syndrome. To analyse postnatal screening to identify risk factors for the subsequent development of type 2 diabetes Study dates All women were diagnosed with gestational diabetes between January 2000 and December 2002	33 (5) Ethnicity, % Caucasian origin: 51 Parity ≥1, % 66 Family history of diabetes, % 47 Pre-pregnancy BMI, kg/m2 25.1 Macrosomia, % 33 Medication during pregnancy, % insulin 75 * The characteristics above are of those who completed the postnatal test Inclusion Criteria All women		-Outcomes: IGT, diabetes -Outcome definitions: WHO 1999 criteria -Timing of postnatal test: 6.4-45.0 weeks -Location of postnatal test (primary/secondary care): unclear -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Source of funding NR	diagnosed with gestational diabetes between January 2000 and December 2002 at the Lausanne University Hospital Exclusion Criteria All patients with pre-existing type 1 or type 2 diabetes				
Full citation Ogonowski,J., Miazgowski,T., The prevalence of 6 weeks postpartum abnormal glucose tolerance in Caucasian women with gestational diabetes, Diabetes Research and Clinical Practice, 84, 239-244, 2009 Ref Id 153592 Country/ies where the study was carried out	Sample size -Number with gestational diabetes: 855 -Number with postnatal test: 318 (37.2%) Characteristics Age in years, mean (SD) 30.96 (0.27) Ethnicity, n(%) Caucasian: 318 (100)	Tests 2-hour 75g OGTT	Methods - All women had 75g OGTT and the following data were collected: age, height, weight, results of the challenge 50g and diagnostic 75g OGTT, and glycated haemoglobin (HbA_{1c}). -Gestational diabetes criteria: Two-step diagnostic procedure using a 50g glucose challenge test and 75g OGTT. Women with a 2-hour glucose level > 200mg/dl (11.1mmol/l) in the challenge test were classified as having gestational diabetes. By the results of diagnostic OGTT,	Results Incidence data Diabetes: 4/318 (1.3%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (exclusion criteria not reported) 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Poland Study type Prospective cohort study Aim of the study To evaluate the incidence of impaired glucose tolerance (IGT), impaired fasting glucose (IFG), and diabetes in 318 Caucasian women with gestational diabetes at 6 weeks' postpartum Study dates All women referred to outpatient clinic for Diabetic Pregnant Women between January 2005 and December 2007 Source of funding NR	Parity NR Family history of diabetes NR Prepregnancy BMI (kg/m2), mean (SD) 24.37 (0.29) Macrosomic infant delivered NR Medication use, % insulin treated 43.3 * The characteristics above are of those who completed the postnatal test Inclusion Criteria - Caucasian women aged > 18		gestational diabetes was diagnosed if either the fasting glucose level was >= 126mg/dl (7.0mmol/l) or the 2-hour glucose concentration was >= 140mg/dl (7.8mmol/l), according to the WHO 1999 criteria -Outcomes: diabetes, IGT, IFG -Outcome definitions: diabetes was diagnosed if either the fasting glucose level was >=126mg/dl (7mmol/l) or the 2-hour glucose concentration was >=200mg/dl (11.1mmol/l), according to the WHO 1999 criteria. IGT was diagnosed if 2-hour glucose was between 140mg/dl and 199mg/dl (7.8 and 11.0mmol/l) and IFG was diagnosed if fasting glucose was between 100mg/dl and 125mg/dl (5.5-6.9mmol/l)-ADA 2003 criteria -Timing of postnatal test: 5-9 weeks (mean 6.0 +/- 0.2 weeks) -Location of postnatal test (primary/secondary care):		verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: nclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported Only data for diabetes was extracted as cut-offs for other outcomes do not match

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	years diagnosed as having glucose intolerance during pregnancy and who were referred to the Outpatient Clinic for Diabetic Pregnant Women in Poland Exclusion Criteria NR		secondary care (women with gestational diabetes were referred to the Outpatient Clinic for Diabetic Pregnant Women in Poland-assuming they returned back here for the follow-up postnatal test) -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		the WHO criteria
Full citation Pallardo,F., Herranz,L., Garcia-Ingelmo,T., Grande,C., Martin-Vaquero,P., Janez,M., Gonzalez,A., Early postpartum metabolic assessment in women with prior gestational diabetes, Diabetes Care, 22, 1053-1058, 1999 Ref Id 153613 Country/ies where the study was	Sample size Number with gestational diabetes: 1425 Number with postnatal test: 788 (55.2%) Characteristics Age in years, mean (SD) 33.1 (11.7) Ethnicity All caucasian women Parity	Tests 75g 2-hour OGTT	Methods -788 women were evaluated 3-6 months after a gestational diabetes pregnancy. A 75g OGTT was performed -Gestational diabetes criteria: 50g oral glucose challenge test at 24-28 weeks' gestation. A positive screen result was defined as 1 hour glucose value >=140mg/dl (7.8mmol/l). Each woman with a positive screen result was given a fasting 3-hour 100g OGTT. The diagnosis of gestational diabetes was made using the criteria of the NDDG.	Results Incidence data Diabetes: 43/788 (5.4%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No, exclusion criteria not reported 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard:

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
carried out Spain Study type Prospective cohort study Aim of the study To present the results of early postnatal metabolic assessment in women with gestational diabetes, to determine predictive factors for subsequent diabetes, and to investigate the association of postnatal glucose tolerance with other components of the metabolic syndrome Study dates All women were seen for the management of gestational diabetes between 1987 and 1997	NR Family history of diabetes, % 50.7 Prepregnancy BMI (kg/m2), mean (SD) 25.9 (16.7) Prior macrosomia, % 10 Medication use, % insulin 49.4 * The characteristics above are of those who completed the postnatal test Inclusion Criteria - In the event of there having been		Gestational diabetes was subclassified according to fasting glucose value as follows: class A1 <105mg/dl (5.8mmol/l); class A2: 105-129mg/dl (5.8-7.2mmol/l) and class B1: >=130mg/dl (7.2mmol/l) -Outcomes: normal, IFG, IGT, IFG IGT, diabetes -Outcome definitions: 1997 ADA criteria. Cut-offs not reported in article but extracted from Conway 1999. Normal: FPG<110mg/dl(6.1mmol/l) and 2hour PG <140mg/dl(7.8mmol/l). IGT: 2hour PG >=140mg/dl(7.8mmol/l) and <200mg/dl(11.1mmol/l). IFG: FPG >=110mg/dl(6.1mmol/l) and <126mg/dl(7mmol/l). Diabetes: FPG >=126mg/dl(7mmol/l)* or 2hour PG >=200mg/dl(11.1mmol/l) *Diagnosis of diabetes based on FPG alone requires that this criterion be confirmed on a second occasion -Timing of postnatal test: 3-6		Yes, whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes (but cut-off levels not stated) 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported Only data for diabetes were extracted as cut-offs for other outcomes do not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Source of funding NR	a subsequent pregnancy complicated by gestational diabetes in the same woman during the years of the study, only the first gestational diabetes pregnancy was considered Exclusion Criteria NR		months postpartum after lactation was concluded -Location of postnatal test (primary/secondary care): secondary care (patients were advised to return to the hospital for testing -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		
Full citation Pallardo,L.F., Herranz,L., Martin-Vaquero,P., Garcia-Ingelmo,T., Grande,C., Janez,M., Impaired fasting glucose and impaired glucose tolerance in women with prior gestational diabetes are associated with a different cardiovascular	Sample size Number with gestational diabetes: 1350 Number with postnatal test: 838 (62%) Characteristics Age in years, mean (SD) 32.4 (4.6) Ethnicity All caucasian	Tests 75g 2-hour OGTT	Methods -838 women with prior gestational diabetes were studied. Postnatal glucose tolerance was classified according to the WHO criteria and postnatal BMI, waist circumference, blood pressure, tryglyceride, cholesterol and high-density lipoprotein (HDL) cholesterol were assessed -Gestational diabetes criteria: Gestational diabetes was diagnosed according to the NDDG criteria after performing a	Results Incidence data Diabetes: 30/838 (3.6%) IFG: 65/838 (7.8%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No, inclusion and exclusion criteria not reported 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
profile, Diabetes Care, 26, 2318-2322, 2003 Ref Id 153614 Country/ies where the study was carried out Spain Study type Prospective cohort study Aim of the study To investigate the association of cardiovascular risk factors to impaired glucose tolerance (IGT) and to impaired fasting glucose (IFG) in women with prior gestational diabetes Study dates Research conducted between 1992 and 2000 Source of funding NR	women Parity, mean (SD) 1.8 (0.9) Family history of diabetes, % NR Prepregnancy BMI (kg/m2), mean (SD) NR Prior macrosomia, % NR Medication use, % insulin 46.1 * The characteristics above are of those who completed the postnatal test Inclusion Criteria		fasting 3-hour 100g OGTT in all pregnant women with a screening test (50g oral glucose challenge) result showing a 1-hour glucose value >=140 mg/dl (7.8mmol/l). Cut-offs for gestational diabetes extracted from reference article; gestational diabetes was diagnosed when two or more glucose values met or exceeded the following thresholds: 5.8mmol/l at fasting, 10.6mmol/l at 1 hour, 9.2mmol/l at 2 hours, 8.1mmol/l at 3 hours. -Outcomes: Normal, IFG, IGT, IFG IGT, Diabetes. -Outcome definitions: WHO criteria with the following modifications: diabetes- fasting glucose >=126mg/dl (7.0mmol/l) or 2-hour glucose >=200mg/dl (11.1mmol/l), IFG- fasting glucose >=110mg/dl (6.1mmol/l) and <126mg/dl (7.0mmol/l) and 2-hour glucose <140mg/dl (7.8mmol/l), IGT- fasting glucose <110mg/dl (6.1mmol/l) and 2-hour glucose >=140mg/dl		selection of the sample receive verification using the reference standard: Yes, whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported Only data for diabetes and IFG were extracted as cut-offs for other outcomes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	NR Exclusion Criteria NR		(7.8mmol/l) and <200mg/dl (11.1mmol/l), IFG plus IGT- fasting glucose >=110mg/dl (6.1mmol/l) and <126mg/dl (7.0mmol/l) and 2-hour glucose >=140mg/dl (7.8mmol/l) and <200mg/dl (11.1mmol/l) and normal- fasting glucose <110mg/dl (6.1mmol/l) and 2-hour glucose <140mg/dl (7.8mmol/l). -Timing of postnatal test: 3-6 months after delivery when lactation was concluded. -Location of postnatal test (primary/secondary care): Unclear -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		do not match the WHO criteria
Full citation Rivero,K., Portal,V.L., Vieira,M., Behle,I., Prevalence of the impaired glucose metabolism and its association	Sample size Number with gestational diabetes: 125 Number with postnatal test: 109 (87.2%)	Tests 75g 2 hour OGTT	Methods -Cohort study of women who gave birth between 1999-2003 and were followed up at the Hospital Padre Jeremias, Cachoeirinha as part of the Day-Hospital Program for women with	Results Incidence data (32 months after delivery) Diabetes: 19/109 (17.4%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
with risk factors for coronary artery disease in women with gestational diabetes, Diabetes Research and Clinical Practice, 79, 433-437, 2008 Ref Id 153690 Country/ies where the study was carried out Brazil Study type Prospective cohort study Aim of the study To investigate the prevalence of type 2 diabetes and IGT and their association with risk factors and inflammatory markers for coronary artery disease among women who had gestational diabetes Study dates	Characteristics Age in years Mean (SD) Normal: 35.19 (7.03) IGT: 35.58 (6.65) Type 2 diabetes: 36.84 (5.69) Ethnicity, n(%) Not reported Parity, Mean (SD) Normal: 3.23 (2.13) IGT: 2.88 (1.49) Type 2 diabetes: 3.94 (3.17) Family history with diabetes, n(%) Not reported Pre-gestational BMI, kg/m2 Normal: 24.60 (4.09) IGT: 26.29 (4.49) Type 2 diabetes: 29.33 (6.03) Current BMI, kg/m2 Normal: 26.29		gestational diabetes -Gestational diabetes criteria: diagnosed with the OGTT: a) with 100g anhydrous glucose (100g-OGTT) according to O'Sullivan et al and as recommended by the ADA in 1997; or b) with 75g anhydrous glucose (75g OGTT) as recommended by the Working Force on Diabetes and Pregnancy and ADA -Outcomes: Diabetes, IGT, Normal -Outcome definitions: Article does not state whether the 1997 or 2003 ADA criteria were used but values match 2003 criteria. Diabetes was defined as FPG >=126mg/dl (7mmol/l) or 2 hour PG>=200mg/dl (11.1mmol/l), IGT as FPG 100-125mg/dl (5.6mmol/l-6.9mmol/l) and/or 2 hour PG 140-199mg/dl (7.8-11.1mmol/l) and Normal as FPG <100mg/dl (5.6mmol/l) and/or 2 hour PG <140mg/dl (7.8mmol/l) -Timing of postnatal test: 6		described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
All women gave birth between 1999 and 2003 Source of funding Not reported	(4.21) IGT: 28.52 (5.09) Type 2 diabetes: 32.24 (6.33) Macrosomic infant delivered Not reported Medication during pregnancy Not reported * The characteristics above are of those who completed the postnatal test Inclusion Criteria -Gestational diabetes women who gave birth during the period 1999-2003 and were followed up at a hospital in Brazil as part of the Day-Hospital Program for women with gestational diabetes*		weeks -Location of postnatal test (primary/secondary care): Hospital Padre Jeremias -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information Only data for diabetes was extracted as cut-offs for other outcomes did not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	*Not explicitly stated as study inclusion criteria in article Exclusion Criteria - Gastrointestinal problems after glucose loading - Withdrawals due to personal questions before screening was completed - Subjects remaining diabetic 6 weeks after delivery - Subjects seen for arterial hypertension without gestational diabetes				
Full citation Schaefer-Graf,U.M., Buchanan,T.A., Xiang,A.H., Peters,R.K., Kjos,S.L., Clinical predictors for a high risk for the development of	Sample size Number with gestational diabetes: 4041 Number with postnatal test: 1636 (40.5%) Characteristics	Tests 75g 2-hour OGTT	Methods -1636 women underwent an OGTT within 1-4 months of delivery. Demographic, historic and antenatal glycaemic parameters and neonatal outcome parameters were tested by univariate and multivariate logistic regression for risk of	Results Incidence data Diabetes: 230/1636 (14.1%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No, inclusion criteria not reported

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
diabetes mellitus in the early puerperium in women with recent gestational diabetes mellitus, American Journal of Obstetrics and Gynecology, 186, 751-756, 2002 Ref Id 153742 Country/ies where the study was carried out USA Study type Retrospective cohort study Aim of the study To identify which maternal, antenatal, or neonatal clinical parameters are predictive for a high risk of diabetes in the puerperium in women with recent gestational diabetes and to calculate the associated diabetes rates and	Age in years, mean (SD) Non-diabetic women: 31.1 (5.8) Women with diabetes: 32.2 (6.0) Ethnicity NR Parity, mean (SD) Non-diabetic women: 1.9 (1.7) Women with diabetes: 2.2 (1.9) Family history of diabetes, % NR Prepregnancy BMI (kg/m2), mean (SD) NR Prior macrosomia >4000g, % Non-diabetic		postnatal diabetes -Gestational diabetes criteria: diagnosed with a 2-step procedure and universal screening policy. If risk factors for gestational diabetes or clinical signs of overt diabetes were present at the initial visit for antenatal care, early screening for gestational diabetes was performed with the use of a 50g 1-hour post-glucose challenge test. Women who were found not to have diabetes were retested between 24 and 28 weeks' gestation. Otherwise, universal screening for gestational diabetes was performed between 24 and 28 weeks' gestation Women with a plasma glucose concentration of 141 to 199mg/dl (7.8mmol/l to 11.1mmol/l) during the 1-hour test were tested for gestational diabetes with a 100g 3-hour OGTT which was interpreted according to the recommendations of the Third International Workshop Conference on gestational diabetes		3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
odds ratios Study dates January 1987-July 1995 Source of funding NR	women: 21.8 Women with diabetes: 32.6 Medication use, % insulin NR * The characteristics above are of those who completed the postnatal test Inclusion Criteria - NR Exclusion Criteria - Incomplete antenatal data		*Eleven women with post-glucose challenge test levels of >=11.1mmol/l or significant glycosuria underwent an initial measurement of FPG levels; an OGTT was only performed if the FPG level was <130mg/dl (7.2mmol/l). Otherwise the diagnosis of gestational diabetes was made on the basis of FPG alone. -Outcomes: diabetes, IFG, IGT -Outcome definitions: during the study period, OGTT results were classified by the NDDG criteria which were current during the study period. For study purposes, diabetes was defined by the new diagnostic criteria of either an overnight FPG level of >=126mg/dl (7mmol/l) or a 2-hour OGTT glucose level of >=200mg/dl (11.1mmol/l) - (ADA 1997 criteria). Criteria used to define IFG/IGT not reported in article -Timing of postnatal test: 1-4		14) Were withdrawals explained: NA Other information -NR: Not reported Only data for diabetes was extracted as this matches WHO.

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
			months after delivery -Location of postnatal test (primary/secondary care): Unclear -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No, although FPG levels were measured before discharge, it is not clear how many women were euglycaemic		
Full citation Schaefer-Graf,U.M., Klavehn,S., Hartmann,R., Kleinwechter,H., Demandt,N., Sorger,M., Kjos,S.L., Vetter,K., bou-Dakn,M., How do we reduce the number of cases of missed postpartum diabetes in women with recent gestational diabetes mellitus?,	Sample size Number with gestational diabetes: 1184 Number with postnatal test: 605 (51.1%) Characteristics Age in years, mean (SD) Of those with a normal OGTT: 32.7 (4.5) Of those with an abnormal OGTT: 32.2 (5.6)	Tests 75g 2-hour OGTT	Methods - In 605 Caucasian women with gestational diabetes, antenatal obstetric and glucose data and the glucose data from postnatal OGTTs performed 13 weeks (median) after delivery were prospectively collected -Gestational diabetes criteria: Fifth International Workshop 2007 criteria for 75g OGTT -Outcomes: diabetes, IFG, IGT -Outcome definitions:	Results Incidence data Diabetes: 33/605 (5.5%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (exclusion criteria not reported) 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Diabetes Care, 32, 1960-1964, 2009 Ref Id 153746 Country/ies where the study was carried out Germany Study type Prospective cohort study Aim of the study To use knowledge of risk factors to develop a model for risk stratification based on the combination of antenatal risk factors that might allow one to distinguish between women with high, intermediate, or low risk for postnatal diabetes within 1 year after gestational diabetes Study dates	Ethnicity, n(%) Caucasian: 605/605 (100%) Parity, mean (SD) Of those with a normal OGTT: 2.2 (1.3) Of those with an abnormal OGTT: 2.5 (1.6) Family history of diabetes (%) Of those with a normal OGTT: 56.6 Of those with an abnormal OGTT: 60.5 Prepregnancy BMI, kg/m2, mean (SD) Of those with a normal OGTT: 25.8 (5.5) Of those with an abnormal OGTT: 28.1 (6.1) Prior		diabetes was diagnosed by a fasting venous plasma glucose >=126mg/dl (7mmol/l) or a 2-hour value >=200mg/dl (11.1mmol/l), IFG by fasting glucose >110mg/dl (6.1mmol/l) and IGT by 2-hour glucose >140 mg/dl (7.7mmol/l) - (similar to ADA 1997). -Timing of postnatal test: 13 weeks (median), within 1 year of delivery -Location of postnatal test (primary/secondary care): NR -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No, at least one glucose profile was performed before discharge only in women with gestational diabetes requiring insulin		selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information -NR: Not reported -Only data for diabetes is extracted as cut-offs for other outcomes in this article do not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Gestational diabetes diagnosed between 1 January 2000 and December 2005 Source of funding NR	macrosomia, % Of those with a normal OGTT: 5.7 Of those with an abnormal OGTT: 7.6 Medication use, % with insulin therapy Unclear reporting (%>100) Inclusion Criteria - Maternal glucose intolerance first diagnosed in pregnancy - Availability of clinical data regarding maternal characteristics, glycaemic data and neonatal parameters - A documented maternal postnatal OGTT				

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	within 1 year of delivery Exclusion Criteria NR				
Full citation Tam,W.H., Yang,X.L., Chan,J.C., Ko,G.T., Tong,P.C., Ma,R.C., Cockram,C.S., Sahota,D., Rogers,M.S., Progression to impaired glucose regulation, diabetes and metabolic syndrome in Chinese women with a past history of gestational diabetes, Diabetes/Metabolism Research Reviews, 23, 485-489, 2007 Ref Id 153847 Country/ies where the study was	Sample size Number with gestational diabetes: 134 Number with postnatal test: 67 (50%) Characteristics Maternal age in years, mean (SD) At index pregnancy: 28.6 (4.3) At 8 year follow-up: 36.9 (4.4) Ethnicity All Chinese women Nulliparity during index pregnancy,n (%) 40 (59.7%)	Tests 2 hour 75g OGTT	Methods -Gestational diabetes criteria: WHO 1999 criteria. On the basis of the 75g OGTT results at the index pregnancy, women were classified as having normal glucose tolerance (FPG <7.0mmol/l and 2 hour plasma glucose <7.8mmol/l) gestational impaired glucose tolerance (FPG <7.0mmol/l and 2 hour plasma glucose ≥7.8-11.1mmol/l) and gestational diabetes (FPG ≥7.0mmol/l and/or 2 hour plasma glucose ≥11.1mmol/l). Cut-off reported for gestational diabetes does not match WHO. -Outcomes: diabetes, IGT, IFG -Outcome definitions: diabetes was defined as FPG ≥7.0mmol/l or 2 hour plasma glucose ≥11.1mmol/l.	Results Incidence data Diabetes: 6/67 (9.0%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
carried out Hong Kong Study type Retrospective cohort study Aim of the study To examine the risk of developing impaired glucose regulation, diabetes and metabolic syndrome in Chinese women with history of gestational diabetes Study dates Subjects were identified from a cohort of women recruited consecutively between 1992 and 1994 Source of funding Supported by Chinese University of Hong Kong Direct Research Grant	Family history of diabetes, n, (%) At index pregnancy: 13 (19.4) At 8 year follow-up: 28 (41.2) BMI, kg/m2 At index pregnancy: 24.8 (3.6) At 8 year follow-up: 24.4 (4.6) Macrosomic infant delivered NR Medication during pregnancy NR * The characteristics above are of those who completed the postnatal test Inclusion Criteria Women with gestational diabetes from the		IGT was defined as FPG <7.0mmol/l and a 2 hour plasma glucose ≥7.8 and <11.1mmol/l. IFG was defined as FPG ≥5.6mmol/l and <7.0mmol/l -Timing of postnatal test: 7-10 years after delivery -Location of postnatal test (primary/secondary care): NR -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information Only data for IGT and diabetes has been extracted as cut-off for IFG given in the article does not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	Prince of Wales Hospital Exclusion Criteria Not reported				
Full citation Xiang,A.H., Kjos,S.L., Takayanagi,M., Trigo,E., Buchanan,T.A., Detailed physiological characterization of the development of type 2 diabetes in Hispanic women with prior gestational diabetes mellitus, Diabetes, 59, 2625-2630, 2010 Ref Id 153940 Country/ies where the study was carried out USA Study type Prospective cohort study Aim of the study	Sample size Number with gestational diabetes: NR Number with postnatal test: 72 Characteristics Maternal age in years, median (IQR) 32.2 (28.2-36.4) Ethnicity All Hispanic women Parity, mean (SD) NR Family history of diabetes, % NR BMI, kg/m2, median	Tests 75g OGTT	Methods -Gestational diabetes criteria: NR -Outcomes: diabetes -Outcome definitions: ADA, diabetes was diagnosed by a fasting glucose ≥7mmol/l or a 2 hour glucose ≥11.1mmol/l -Timing of postnatal test: 15-30 months after delivery -Location of postnatal test (primary/secondary care): Los Angeles County Women's Hospital -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No	Results Incidence data During a median follow-up of 72 months (range:12-142months) Diabetes: 31/72 (43%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
To identify physiological and clinical variables associated with development of type 2 diabetes up to 12 years after pregnancies complicated by gestational diabetes Study dates All women were referred to the hospital for management of gestational diabetes between August 1993 and March 1995 Source of funding Grants from the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, National Center for Research Resources and a Distinguished Clinical Scientist	(interquartile range) 30.7 (27.8-32.8) Macrosomia (%) NR Medication during pregnancy, % insulin None, as inclusion criteria was no current or prior insulin therapy Inclusion Criteria - Gestational age between 28 and 34 weeks - No current or prior insulin therapy - All fasting serum glucose concentrations <7.2mmol/l during pregnancy - Otherwise uncomplicated singleton pregnancy - Both parents and at least three of four				described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: Yes Other information NR: Not reported Only diabetes data has been extracted as cut-offs for IFG and IGT do not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Award from the American Diabetes Association	grandparents were from Mexico, Guatemala or El Salvador Exclusion Criteria NR				
Full citation Kerimoglu,O.S., Yalvac,S., Karcaaltincaba,D., Kandemir,O., Altinbas,S.K., Dede,H., Early post-partum diabetes mellitus screening rates in patients with history of gestational diabetes, Archives of Gynecology and Obstetrics, 282, 613-616, 2010 Ref Id 154131 Country/ies where the study was carried out Turkey Study type	Sample size Number with gestational diabetes: 78 Number with postnatal test: 37/78 (47%) Characteristics Maternal age in years, median (IQR) Of those evaluated with 75g OGTT: 37 (5.8) Of those evaluated with FPG: 35 (4) Ethnicity, n (%) NR	Tests 75g OGTT FPG only: 27/78 (34.6%) OGTT: 10/78 (12.8%)	Methods -The study included 78 women diagnosed and treated for gestational diabetes. They were evaluated whether or not they were screened with 75g OGTT or FPG at 6-12 weeks postpartum. The rates of diabetes and impaired glucose tolerance were determined -Gestational diabetes criteria: NDDG criteria. Two-step process- 50g 1-hour glucose challenge test and then a 100g 3-hour diagnostic OGTT if glucose challenge test result >=140mg/dl(7.8mmol/l). Gestational diabetes was diagnosed when two or more glucose values during the diagnostic OGTT met or exceeded the criteria for a	Results Incidence data OGTT Diabetes: 5/10 (50%) FPG Diabetes: 2/27 (7.4%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (exclusion criteria not reported) 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: No (only 10/78 completed OGTT) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Retrospective cohort study Aim of the study To investigate the rate of gestational diabetes women who received screening only by FPG measurement or OGTT and the prevalence of diabetes detected by these screening tests early in the postnatal period Study dates All women with gestational diabetes diagnosed and hospitalised for glucose regulation between 2005-2007 Source of funding NR	Primiparous, n (%) Of those evaluated with 75g OGTT: 1 (10) Of those evaluated with FPG: 5 (17.9) Family history of diabetes in first degree relatives, n (%) Of those evaluated with 75g OGTT: 9 (90) Of those evaluated with FPG: 16 (59.3) BMI (kg/m2), mean (SD) NR Medication use during pregnancy, n (%) Diet only Of those evaluated with 75g OGTT: - Of those		positive test - plasma glucose thresholds: fasting 95mg/dl (5.3mmol/l), 1 hour 180 mg/dl (10mmol/l), 2 hours 155mg/dl (8.6mmol/l), 3 hours 140mg/dl (7.8mmol/l) - these cut-offs do not match the cut-offs in the NDDG reference article -Outcomes: Diabetes, IGT, IFG -Outcome definitions: ADA criteria - diabetes: 2-hour postload glucose >=200mg/dl (11.1mmol/l) or FPG>=126mg/dl(7mmol/l), IGT: 2-hour postload glucose 140-199mg/dl (7.8-11.1mmol/l), IFG: FPG 100-125mg/dl (5.6-6.9mmol/l). Article does not report whether 1997 or 2003 criteria were used but cut-offs match the 2003 criteria -Timing of postnatal test: 6-12 weeks -Location of postnatal test (primary/secondary care): NR -Did study document a return to euglycaemia in the		independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported Only data for diabetes was extracted as cut-offs for other outcomes in this article do not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	evaluated with FPG: 13 (48.1) Insulin added Of those evaluated with 75g OGTT: 10 (100) Of those evaluated with FPG: 14 (51.9) Macrosomic infant delivered NR * The characteristics above are of those who completed the postnatal test Inclusion Criteria -Gestational diabetes patients who were hospitalised during their pregnancy because they were better informed about their disease and risk of development of		immediate days following delivery and before discharge: No		

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	diabetes in the future Exclusion Criteria NR				
Full citation Retnakaran,R., Qi,Y., Sermer,M., Connelly,P.W., Zinman,B., Hanley,A.J., Comparison of National Diabetes Data Group and American Diabetes Association diagnostic criteria for gestational diabetes in their identification of postpartum risk of glucose intolerance, Diabetes Research and Clinical Practice, 85, 40-46, 2009 Ref Id 154244 Country/ies where the study was carried out	Sample size 284 women with GDM/IGT underwent postnatal test. Characteristics Antenatally Maternal age in years antenatally, mean (SD) In those with IGT by ADA: 34 (4.3) In those with gestational diabetes by ADA: 34.9 (4.3) Ethnicity % White In those with IGT by ADA only: 85.7	Tests 2 hour 75g OGTT	Methods -Gestational diabetes criteria: Based on 4 blood glucose values obtained during the 3 hour 100g OGTT (fasting, 1,2,3 hour glucose), subjects were classified as either having gestational diabetes (defined by two or more values above criterion thresholds), IGT (defined by only one value above criterion thresholds) or normal glucose tolerance. The ADA thresholds are i) fasting <5.3mmol/l, ii) 1 hour glucose <10.0mmol/l, iii) 2 hour glucose <8.6mmol/l iv) 3 hour glucose <7.8mmol/l. The NDDG thresholds are i) fasting <5.8mmol/l ii) 1 hour <10.6mmol/l, iii) 2 hour glucose <9.2mmol/l, iv) 3 hour glucose <8.1mmol/l -Outcomes: IFG, IGT, IFG and IGT, diabetes	Results Incidence data IFG: 1.1% (3*/284) Diabetes: 3.2% (9*/284) *Calculated by NCC-WCH technical team	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Canada Study type Prospective cohort study Aim of the study To systematically compare NDDG and ADA criteria in their identification of postnatal risk of glucose intolerance in a well-characterised cohort of women undergoing metabolic characterisation in pregnancy and in the postnatal period. Study dates Not reported Source of funding Canadian Institutes of Health Research (CIHR) operating grants	In those with gestational diabetes by ADA only: 74.5 % Asian In those with IGT by ADA only: 6.1 In those with gestational diabetes by ADA only: 17.6 % Other In those with IGT by ADA only: 8.2 In those with gestational diabetes by ADA only: 7.8 Parity, mean (SD) NR Family history of diabetes, % In those with IGT by ADA only: 55.1 In those with gestational diabetes by ADA only: 49.0 Pre-pregnancy		-Outcome definitions: Cut-offs not reported in article but extracted from a reference article. Diabetes defined as FPG >=7.0mmol/l or 2 hour glucose >=11.1mmol/l. IGT defined by FPG <6.1mmol/l and 2 hour glucose 7.8-11.0mmol/l inclusive. IFG defined as FPG 6.1-6.9mmol/l inclusive, with 2 hour <7.8mmol/l. Combined IFG/IGT defined as FPG 6.1-6.9mmol/l inclusive and 2 hour 7.8-11.0mmol/l inclusive -Timing of postnatal test: 3 months postpartum -Location of postnatal test (primary/secondary care): Unclear -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information Only data for IFG and diabetes has been extracted as cut-offs for other outcomes do not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	BMI, kg/m2 In those with IGT by ADA only: 25.7 (23-30) In those with gestational diabetes by ADA only: 24.0 (22-28) BMI at 3 months postpartum In those with IGT by ADA only: 28.1 (25-31) In those with gestational diabetes by ADA only: 26.4 (23-30) Medication use during pregnancy, n (%) NR Macrosomic infant delivered NR * The characteristics above are of those who completed the postnatal test				

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	Inclusion Criteria NR Exclusion Criteria NR				
Full citation Stasenko,M., Cheng,Y.W., McLean,T., Jelin,A.C., Rand,L., Caughey,A.B., Postpartum follow-up for women with gestational diabetes mellitus, American Journal of Perinatology, 27, 737-742, 2010 Ref Id 154287 Country/ies where the study was carried out USA Study type Retrospective cohort study Aim of the study To evaluate the frequency of postnatal follow-	Sample size Number with gestational diabetes: 745 Number with postnatal test: 251 (33.7%) Characteristics Maternal age in years, n (%) <35: 133/251 (53) >/=35: 118/251 (47) Ethnicity, n (%) White: 66/246 (27) African-American: 16/246 (7) Latina: 18/246 (7) Asian: 146/246 (59) Parity, n (%)	Tests FPG or 2-hour 75g OGTT	Methods - A retrospective cohort study of women with gestational diabetes. Primary outcome was either a FPG or a 2-hour OGTT, both measured at </=6 months postpartum. Chi-square test and multivariable logistic regression analysis were used for statistical comparisons, and statistical significance was indicated by p<0.05 and 95%CIs - Gestational diabetes criteria: Carpenter-Coustan crtieria, 3-hour OGTT: two elevated values on a 3-hour glucose tolerance test utilizing thresholds of 95mg/dl (5.3mmol/l) fasting, 180mg/dl (10mmol/l) at 1 hour, 155mg/dl (8.6mmol/l) at 2 hours and 140mg/dl (7.8mmol/l) at 3 hours post-glucose load -Outcomes: IGT, type 2	Results Incidence data Diabetes: 5/251 (2.0%)* *Elevated FPG or OGTT consistent with type 2 diabetes	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (inclusion and exclusion criteria not reported) 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Unclear(women were considered tested if she had a documented FPG or 2-hour OGTT, not clear how many had OGTT) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
up screening of women with gestational diabetes in a racially/ethnically and socioeconomically diverse population, to identify groups with particularly low follow-up frequency, to provide tailored public health measures to improve care and to elucidate which strata are at a greater risk for developing type 2 diabetes Study dates All women with gestational diabetes delivered between 2002 and 2008 Source of funding One author was supported by a Robert Wood Johnson	Multiparous: 117/251 (47) Nulliparous: 134/251 (53) Family history of diabetes NR Maternal BMI, n (%) <25: 96/199 (48) >/=25: 103/199 (52) Macrosomic infant NR Medication use during pregnancy, n (%) insulin 190/251 (76) * The characteristics above are of those who completed the postnatal test Inclusion Criteria		diabetes -Outcome definitions: 1) IGT: FPG >=95mg/ml (5.3mmol/l) or 2-hour OGTT >=140mg/ml (7.8mmol/l) 2) Type 2 diabetes: FPG >=126mg/ml (7mmol/l) or 2 hour OGTT >=200mg/ml (11.1mmol/l) - Name of criteria not reported. -Timing of postnatal test: <=6 months -Location of postnatal test (primary/secondary care): secondary care (assuming women returned to the hospital that issued a laboratory slip to obtain postnatal testing) -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported Only data for diabetes was extracted as the cut-offs for other outcomes do not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Physician Faculty Scholar Grant	NR Exclusion Criteria NR				
Full citation Ekelund,M., Shaat,N., Almgren,P., Groop,L., Berntorp,K., Prediction of postpartum diabetes in women with gestational diabetes mellitus, Diabetologia, 53, 452-457, 2010 Ref Id 154355 Country/ies where the study was carried out Sweden Study type Prospective cohort study Aim of the study To study the incidence of postnatal diabetes after gestational	Sample size Number with gestational diabetes: 188, 174 had repeated OGTT at inclusion Number with postnatal test: At 1 year 123 out of 174, At 2 year 85 out of remaining 159, at 5 years 112 out of remaining 152 Characteristics Maternal age at delivery in years, mean (SD) In those with NGT at 5 years postpartum: 31.0 (4.6) In those with IGT-IFG at 5 years postpartum: 32.0 (5.9) In those with diabetes at 5	Tests 75g OGTT	Methods -Gestational diabetes criteria: 75g OGTT, a 2 hour capillary blood glucose ≥9mmol/l was defined as the diagnostic threshold of gestational diabetes -Outcomes: IFG, IGT, diabetes -Outcome definitions: WHO 1999 criteria. Cut-offs not reported in article -Timing of postnatal test: 1,2,5 years postpartum -Location of postnatal test (primary/secondary care): Department of Endocrinology -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No	Results At 1 year Diabetes: 12.2% (15/123) At 2 years Diabetes: 8.2% (7/85) At 5 years Diabetes: 12.5% (14/112) IGT: 24.1% (27/112) IFG: 3.6% (4/112)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes at 1 year and then those that tested negative underwent OGTT 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
diabetes and to investigate biochemical and clinical predictors of postnatal diabetes Study dates All women diagnosed with gestational diabetes were referred for follow-up during pregnancy between 1996 and 1999 Source of funding Supported by the Zoégas Foundation, Lundström Foundation, Research Funds of Malmö University Hospital and by grants from County of Skåne	years postpartum: 31.6 (5.8) Ethnicity, Swedish origin, n(%) In those with NGT at 5 years postpartum: 41 (59) In those with IGT-IFG at 5 years postpartum: 8 (26) In those with diabetes at 5 years postpartum: 18 (42) Number with previous pregnancies, n (%) In those with NGT at 5 years postpartum: 33 (47) In those with IGT-IFG at 5 years postpartum: 24 (80) In those with diabetes at 5 years postpartum: 31 (72)				described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported NGT: normal glucose tolerance

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	Family history of diabetes, % In those with NGT at 5 years postpartum: 34 (49) In those with IGT-IFG at 5 years postpartum: 17 (55) In those with diabetes at 5 years postpartum: 30 (70) BMI during pregnancy, kg/m2, median (range) In those with NGT at 5 years postpartum: 27.0 (25.8-29.9) In those with IGT-IFG at 5 years postpartum: 29.3 (26.2-32.0) In those with diabetes at 5 years postpartum: 30.9 (27.1-32.9) Macrosomia (%) NR				

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	Medication during pregnancy, n (%) insulin In those with NGT at 5 years postpartum: 1 (1) In those with IGT-IFG at 5 years postpartum: 5 (16) In those with diabetes at 5 years postpartum: 13 (30) * The characteristics above are of those who completed the postnatal test Inclusion Criteria -All women diagnosed with gestational diabetes Exclusion Criteria - Those subjects for which a repeat OGTT at study start could not be				

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	performed				
Full citation Lawrence,J.M., Black,M.H., Hsu,J.W., Chen,W., Sacks,D.A., Prevalence and timing of postpartum glucose testing and sustained glucose dysregulation after gestational diabetes mellitus, Diabetes Care, 33, 569-576, 2010 Ref Id 154373 Country/ies where the study was carried out USA Study type Retrospective cohort study Aim of the study To estimate the prevalence of postnatal glucose testing within 6 months of	Sample size Number with gestational diabetes: 11825 Number with postnatal test: 7 days to <6 weeks: n=2596 6-12 weeks: 2728 >12 weeks to 6 months: 533 Characteristics Age in years (%) 13-19: 47 (0.8) 20-24: 354 (6) 25-29: 1349 (23) 30-34: 2032 (34) 35-39: 1643 (28) >=40: 514 (9) Ethnicity, n (%) Hispanic: 3139 (53) Black: 219 (4) Asian/Pacific Islander: 1333 (22) Other/unknown: 64 (1)	Tests FPG only: 4698 (79.1%) OGTT: 1081 (18.2%) FPG and OGTT: 160 (2.7%)	Methods - A retrospective study of 11825 women with gestational diabetes. Postpartum tests included the 75g 2-hour OGTT or FPG within 6 months of delivery. Postpartum test results were categorised as normal, IFG, and/or IGT and 'provisionally diabetic' -Gestational diabetes criteria: ADA criteria-100g 3-hour OGTT identified women who had gestational diabetes based on at least two abnormal plasma glucose measurements greater than or equal to the Carpenter and Coustan threshold values recommended by the ADA - fasting 95mg/dl (5.3mmol/l), 1 hour 180mg/dl (10mmol/l), 2 hours 155mg/dl (8.6mmol/l), 3 hours 140mg/dl (7.8mmol/l) -Outcomes: Normal, IFG, IGT, provisional diabetes -Outcome definitions: The ADA criteria were used to	Results Incidence data (based on FPG or OGTT*) 7 days to <6 weeks, n (%) Provisional diabetes: 16 (0.6) 6-12 weeks, n (%) Provisional diabetes: 27 (1.0) >12 weeks to 6 months, n (%) Provisional diabetes: 23 (4.3) *only 18.2% of all subjects had OGTT	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: No (only 18.2%) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
pregnancies complicated by gestational diabetes, to assess factors associated with testing and timing of testing after delivery, and report test results among tested women Study dates All women identified as having gestational diabetes from 1 January 1999 and 31 December 2006 Source of funding Supported by the American Diabetes Association with additional support from Kaiser Permanente Southern California (KPSC) Direct Community Benefit funds	Non-Hispanic white: 1184 (20) Parity, n (%) 0: 2213 (37) 1: 1866 (31) >=2: 1860 (31) Unknown: 0 (0) Family history of diabetes NR BMI NR Macrosomic infant delivered NR Medication use (gestational diabetes treatment), n (%) None: 4530 (76) Insulin (+/-oral agents): 1236 (21) Oral agents only: 173 (3) * The		classify women with an FPG (whether alone or as part of a 75g OGTT) <100mg/dl (5.6mmol/l) as normal, 100-125mg/dl (5.6-6.9mmol/l) as IFG, and >=126mg/dl (7mmol/l) as having a provisional diagnosis of diabetes. Categories based on the glucose concentration 2 hours after a 75g post-glucose load were as follows: <140mg/dl (7.8mmol/l) normal, 140-199mg/dl (7.8-11.1mmol/l) IGT, and >=200mg/dl (11.1mmol/l) provisionally diabetic. Women with IFG and/or IGT were combined into one category. Article does not state whether 1997 or 2003 criteria was used but cut-offs match 2003 criteria. -Timing of postnatal test: 7 days postpartum-6 weeks postpartum (early testing window), 6-12 weeks postpartum (ADA recommended testing window), after 12 weeks-6 months postpartum (late testing window) -Location of postnatal test		permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: No 14) Were withdrawals explained: NA Other information NR: Not reported Only data for diabetes has been extracted as cut-offs for other outcomes do not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	characteristics above are of those who completed the postnatal test Inclusion Criteria - Women who had one or more singleton births at >=20 weeks' gestation in KPSC hospitals, who were identified as having gestational diabetes using the 100-g OGTT from 1 January 1999 through to 31 December 2006, and who remained KPSC members for at least 6 months postpartum Exclusion Criteria - Women with evidence of diabetes before pregnancy		(primary/secondary care): NR -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		
Full citation	Sample size	Tests	Methods	Results	Limitations

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Kim,C., Herman,W.H., Cheung,N.W., Gunderson,E.P., Richardson,C., Comparison of hemoglobin A_{1c} with fasting plasma glucose and 2-h postchallenge glucose for risk stratification among women with recent gestational diabetes mellitus, Diabetes Care, 34, 1949-1951, 2011 Ref Id 157584 Country/ies where the study was carried out USA Study type Prospective cohort study Aim of the study To examine the agreement between A_{1C}, FPG and 2 hour glucose among	54 women with gestational diabetes underwent postnatal test Characteristics Maternal age (years) 36 +/- 4 Ethnicity,% Non-Hispanic white: 73 Asian: 11 African American: 11 Parity Not reported BMI, kg/m2 30.6 +/-7.0 Family history of diabetes Not reported Medication during pregnancy Not reported Inclusion Criteria - Physician confirmed gestational diabetes diagnosis within the past 3 years	2 hour 75g OGTT	-Study assessed the association of A_{1C} >=5.7% with FPG >=100mg/dl(5.6mmol/l) and 2 hour glucose >=140mg/dl(7.8mmol/l) among 54 women with histories of gestational diabetes between 6 weeks and 36 months postpartum -Gestational diabetes criteria: Physician confirmed gestational diabetes diagnosis (details not reported) -Outcomes: Diabetes, IFG, IGT -Outcome definitions: Diabetes defined as FPG >=126mg/dl(7mmol/l) and/or 2 hour glucose >=200mg/dl(11.1mmol/l). FPG >=100mg/dl(5.6mmol/l) as consistent with IFG or diabetes, 2 hour values >=140mg/dl(7.8mmol/l) as consistent with IGT or diabetes and A_{1C} >=5.7% as consistent with increased risk of diabetes. -Timing of postnatal test: 6 weeks to 36 months	Incidence data Diabetes: 5/54 (9.3%) A_{1C} >=5.7: 25/54 (46.3%)	NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No, exclusion criteria not reported 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
women with recent gestational diabetes Study dates Not reported Source of funding Supported by National Institutes of Health grants; the Chemistry Core of the Michigan Diabetes Research and Training Center funded by the National Institute of Diabetes and Digestive and Kidney Diseases; a Robert Wood Johnson Physician Faculty Scholars Program Award; and a Family Medicine Research Pilot Funds Grant	- No pre-existing diabetes diagnosis - Enrolment at >=6 weeks after delivery - Age >=18 years - <150 minutes of self-reported physical activity per week and no contraindications to walking - Fluency in English - Working email address - Lack of current pregnancy, confirmed by a study urine pregnancy test Exclusion Criteria Not reported		postpartum -Location of postnatal test (primary/secondary care): Not reported -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information Only data for diabetes was extracted as cut-offs for other outcomes do not match the WHO criteria
Full citation Krishnaveni,G.V., Hill,J.C.,	Sample size Number with gestational	Tests 2 hour 75g	Methods Gestational diabetes criteria: 100g 3 hour OGTT,	Results Incidence data	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Veena,S.R., Geetha,S., Jayakumar,M.N., Karat,C.L., Fall,C.H., Gestational diabetes and the incidence of diabetes in the 5 years following the index pregnancy in South Indian women, Diabetes Research and Clinical Practice, 78, 398-404, 2007 Ref Id 157623 Country/ies where the study was carried out India Study type Prospective cohort study Aim of the study To examine the incidence of diabetes and the factors associated with this in a cohort of South Indian women 5 years after they were examined for	diabetes: 41 Number with postnatal test: 35 (85%) Characteristics Maternal age in years, mean (range) In women with normal glucose tolerance: 32.2 (28.0, 36.0) In women with IGT/IFG: 34.0 (30.0, 38.0) In women with diabetes: 35.5 (29.5, 38.5) Ethnicity,% NR Parity > 2, n (%) In women with normal glucose tolerance: 1 (9) In women with IGT/IFG: 2 (18) In women with diabetes: 3 (23) BMI, kg/m2 In women with normal glucose	OGTT	gestational diabetes was diagnosed using the Carpenter Coustan criteria -Outcomes: Diabetes, IGT, IFG -Outcome definitions: Diabetes was defined as a fasting glucose >=7.0 and/or 2 hour glucose >=11.1mmol/l. Women were also classified as having diabetes if they had been diagnosed by a doctor as having diabetes since the index pregnancy. IGT was defined as a fasting glucose concentration <7.0mmol/l and 2 hour glucose >=7.8mmol/l but <11.1mmol/l. IFG was defined as a fasting glucose value >=6.1mmol/l and <7.0mmol/l (WHO 1999) -Timing of postnatal test: 5 years -Location of postnatal test (primary/secondary care): unclear -Did study document a return to euglycaemia in the immediate days following	IGT/IFG: 11/35 (31%) Diabetes: 13/35 (37%)	diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
gestational diabetes Study dates Gestational diabetes was diagnosed between 1997 and 1998 Source of funding The Parthenon Trust, Switzerland, the Wellcome Trust UK and the Medical Research Council, UK	tolerance: 23.6 (4.4) In women with IGT/IFG: 26.1 (3.0) In women with diabetes: 26.7 (4.6) Family history of diabetes In women with normal glucose tolerance: 5 (46%) In women with IGT/IFG: 3 (27%) In women with diabetes: 12 (92%) Medication during pregnancy, n (%) In women with normal glucose tolerance: 0 (0) In women with IGT/IFG: 3 (27.3) In women with diabetes: 4 (30.8) Inclusion Criteria All willing, non-pregnant women, who had not been		delivery and before discharge: No		results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	pregnant within the previous 6 months (previous gestational diabetes pregnancy). Examination of these women were based on the follow-up of their offspring Exclusion Criteria 7 children (and therefore their mothers) were excluded after birth due to medical reasons				
Full citation Katon,J., Reiber,G., Williams,M.A., Yanez,D., Miller,E., Hemoglobin a_{1c} and postpartum abnormal glucose tolerance among women with gestational diabetes mellitus, Obstetrics and Gynecology, 119,	Sample size Number with gestational diabetes: 536 Number with postnatal test: 277 (52%) Characteristics Maternal age at gestational diabetes diganosis in	Tests 75g 2-hour OGTT	Methods -Women with singleton pregnancies treated for gestational diabetes at a large diabetes and pregnancy programme in North Carolina who completed a postnatal 2-hour OGTT were included in this retrospective cohort study. Clinical information was abstracted from medical records -Gestational diabetes	Results Incidence data, n (%) Diabetes: 15/277 (5%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
566-574, 2012 Ref Id 157640 Country/ies where the study was carried out USA Study type Retrospective cohort study Aim of the study To analyse the association of HbA_{1c} at gestational diabetes diagnosis with postnatal abnormal glucose in a cohort of women with gestational diabetes Study dates All women delivered between November 15, 2000 and April 15, 2010 Source of funding One author was supported by a grant from: the	years, mean (SD) 31 (5.2) Ethnicity, n (%) White: 104 (38) African American: 51 (18) Hispanic: 88 (32) Asian Indian: 27 (10) Other: 7 (2) Parity, n (%) Nulliparous: 121 (44) Family history of diabetes NR Prepregnancy BMI, kg/m2, n (%) <25: 91 (33) 25-29.9: 84 (30) >/=30: 102 (37) Macrosomic infant delivered NR Medication use		criteria: Two-step process: 50g oral challenge and then 3-hour 100g OGTT if abnormal (NDDG criteria) -Outcomes: IFG (with or without impaired glucose tolerance), IGT(with or without impaired fasting glucose) and any postpartum abnormal glucose including type 2 diabetes -Outcome definitions: ADA criteria 1) Normal glucose: FPG <100mg/dl (5.6mmol/l), 2-hour plasma glucose <140mg/dl (7.8mmol/l) 2) IFG: FPG>= 100mg/dl (5.6mmol/l) and < 126mg/dl (7mmol/l) 3) IGT: 2-hour plasma glucose >=140mg/dl (7.8mmol/l) and < than 200mg/dl (11.1mmol/l) 4) Type 2 diabetes: FPG >= 126mg/dl (7mmol/l) or 2 hour plasma glucose >=200mg/dl (11.1mmol/l). -Timing of postnatal test: Median-7.9 weeks, IQR-6.6-9.4, Range-3-111 -Location of postnatal test (primary/secondary care):		that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health; the Seattle chapter of Achievement Rewards for College Scientists; and the Samuel and Althea Stroum Foundation. The study was also funded by a grant from the University of Washington Department of Epidemiology	(gestational diabetes), n (%) Diet only: 56 (20) Glyburide: 46 (17) Insulin: 162 (58) Metformin: 3 (1) Other: 10 (4) * The characteristics above are of those who completed the postnatal test Inclusion Criteria - Women treated for gestational diabetes who delivered a live singleton neonate between November 15, 2000 and April 15, 2010 - Diagnosed with gestational diabetes at 24 weeks' gestation or greater by a 3-hour 100g OGTT, a glucose challenge test 200mg/dl or		NR -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		Only data for diabetes was extracted as cut-offs for other outcomes do not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	higher, or a random blood glucose 160mg/dl (8.9mmol/l) or higher and completion of a postnatal 2-hour 75g OGTT Exclusion Criteria - Established type 1 or type 2 diabetes - Gestational diabetes diagnosis at less than 24 weeks' gestation, untreated endocrinopathies (hyperadrenalism, hypoadrenalism, hyperthyroidism, hypothyroidism and acromegaly), haemoglobin variants (HbS, HbC, HbF, HbE) or conditions (uraemia, thalassaemia) that impair interpretation of HbA_{1c}				

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	-First HbA_{1c} measurement more than 4 weeks after the initial visit to the diabetes and pregnancy programme -Use of medications at the time of postnatal OGTT that affect glucose tolerance (metformin, glyburide, steroids, hydrochlorothiazide) -Pregnant at the time of the postnatal OGTT				
Full citation Hossein-nezhad,A., Mirzaei,K., Maghbooli,Z., Larijani,B., Maternal glycemic status in GDM patients after	Sample size Number with gestational diabetes: 114 Number with postnatal test: 98 (86%) Characteristics	Tests 2 hour 75g OGTT	Methods -Gestational diabetes criteria: 2 step procedure using a 50g glucose challenge test and a 75g OGTT. All women with plasma glucose values >=130mg/dl were given an 100g 3 hour glucose	Results Diabetes: 8.1% (8/98)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
delivery, Iranian Journal of Diabetes and Lipid Disorders, 8, 95-104, 2009 Ref Id 157679 Country/ies where the study was carried out Iran Study type Prospective cohort study Aim of the study To examine the association between gestational diabetes and susceptibility to type 2 diabetes and impaired glucose tolerance after pregnancy Study dates NR Source of funding Grant from Endocrinology and Metabolism	Maternal age at gestational diabetes diganosis in years, mean (SD) 29 (6) Ethnicity, n (%) Not reported Parity 1 (3) Family history of diabetes, % 33.3 Prepregnancy BMI in kg/m2, mean (SD) 27.4 (4.3) Macrosomic infant delivered, % 25.4 Medication use (gestational diabetes), (%) 16.3%		tolerance test to diagnose gestational glucose intolerance using the Carpenter Coustan criteria -Outcomes: IFG, IGT, diabetes -Outcome definitions: ADA criteria. Diabetes was diagnosed if the fasting blood glucose was >=7mmol/l. IGT was diagnosed if the 2 hour postprandial glucose was between 7.8 and 11.0mmol/l and IFG was diagnosed if fasting glucose was between 5.5 and 6.9mmol/l -Timing of postnatal test: 6-12 weeks -Location of postnatal test (primary/secondary care): NR -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Research Center	Inclusion Criteria Women consecutively referred to 5 university educational hospitals in Tehran, Iran for antenatal care Exclusion Criteria NR				or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information Only data for diabetes has been extracted as cut-offs for other outcomes do not match the WHO criteria
Full citation Anderberg,E., Landin-Olsson,M., Kalen,J., Frid,A., Ursing,D., Berntorp,K., Prevalence of impaired glucose tolerance and diabetes after gestational diabetes mellitus comparing different cut-off criteria for abnormal glucose tolerance during pregnancy, Acta Obstetricia et Gynecologica Scandinavica, 90, 1252-1258, 2011	Sample size Number with gestational diabetes: 298 Number with postnatal test: 160/298 (54%) Characteristics Age at delivery in years, mean (SD) 33.1 (4.9) Ethnicity, n(%) Swedish origin: 92/160 (58) European origin except Swedish: 25/160 (16) Non-European origin: 43/160	Tests 75g OGTT	Methods -Gestational diabetes criteria: 75g OGTT, 2-hour capillary blood glucose concentration >=9.0mmol/l (plasma glucose >=10.0mmol/l)- The Diabetes Pregnancy Study Group of the European Association for the Study of Diabetes (EASD) -Outcomes: Diabetes, IGT -Outcome definitions: WHO 1999 criteria. Diabetes- FPG >=7mmol/l (126mg/dl) and/or 2-hour PG>=11.1mmol/l (200mg/dl). IGT- FPG<7mmol/l (126mg/dl) and 2-hour PG 7.8-11.0mmol/l (140-199mg/dl)	Results Incidence data Diabetes: 17/160 (11%) IGT: 38/160 (24%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (inclusion and exclusion criteria not reported) 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample)

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Ref Id 157717 Country/ies where the study was carried out Sweden Study type Prospective cohort study Aim of the study To evaluate the frequency of abnormal glucose tolerance postnatal when lowering the cut-off level for gestational diabetes to include milder forms of IGT during pregnancy, and to identify a target group for primary diabetes prevention Study dates All women delivered between 2003 and 2005 Source of funding This study was	(27) Parity, n (%) Nulliparous: 65 (42) First degree relative with diabetes, n (%) 61 (42) BMI NR Macrosomic infant delivered NR Medication during pregnancy NR * The characteristics above are of those who completed the postnatal test Inclusion Criteria NR Exclusion Criteria Subjects already		-Timing of postnatal test: 1-2 years after delivery -Location of postnatal test (primary/secondary care): secondary care (diabetes care unit in a hospital) -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
supported by the Research Funds of Malmo and Lund University Hospitals, and the Foundations of the County of Skane	diagnosed with diabetes				
Full citation Saucedo,R., Zarate,A., Basurto,L., Hernandez,M., Puello,E., Campos,S., Moreno,E., Women with gestational diabetes develop glucose intolerance with high frequency within one year postpartum, Gynecologic and Obstetric Investigation, 73, 58-62, 2012 Ref Id 157755 Country/ies where the study was carried out Mexico	Sample size Number with gestational diabetes: 100 Number with postnatal test: 52 (52%) Characteristics Maternal age (years) Normal : 26.6 ± 1.5 IFG/IGT : 31.5 ± 3.2 Diabetes : 33.5 ± 4.7 Race/ethnicity NR Parity % Normal : Nulliparous 34.0, 1 pregnancy = 66.0, >1 pregnancy = 0	Tests 75g 2-hour OGTT	Methods -Gestational diabetes criteria: Women were screened for gestational diabetes using a 2-hour 75g OGTT at 24-28 weeks' gestation and cutoff values of >95.0mg/dl (5.3mmol/l) fasting, >180mg/dl (10mmol/l) at 1 hour and >155.0mg/dl (8.6mmol/l) at 2 hours - ADA -Outcomes: IFG, IGT or diabetes -Outcome definitions: The article does not report whether the 1997 or 2003 ADA criteria were used but values match 2003 criteria. Normal glucose tolerance defined as FPG <100 mg/dl (5.6mmol/l) and a 2-hour plasma glucose value <140 mg/dl (7.8mmol/l)	Results Incidence data At 6 weeks after delivery Diabetes : 9/52 (17.3%) At 6 months after delivery Diabetes : 17/52 (32.7%) At 1 year after delivery Diabetes : 25/52 (48.1%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes, whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test? (that is, the index test did not form part of the reference standard): No

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Study type Prospective cohort study Aim of the study To examine the incidence of postnatal glucose intolerance in women with gestational diabetes and to assess their body weight, cholesterol and triglyceride concentrations after delivery Study dates July 2007 to May 2009 Source of funding Grants from IMSS and CONACYT	IFG/IGT : Nulliparous 19.0, 1 pregnancy = 28.6, >1 pregnancy = 52.4 Diabetes : Nulliparous 14.2, 1 pregnancy = 17.9, >1 pregnancy = 67.9 Family history of diabetes (%) Normal : 33.3 IFG/IGT : 66.6 Diabetes : 70.4 BMI : Normal : 28.2 ± 4.5 IFG/IGT : 31.3 ± 4.7 Diabetes : 32.8 ± 4.5 Macrosomic infant delivered NR Insulin use during pregnancy (%) Normal : 0 IFG/IGT : 47.6 Diabetes : 75.0 Inclusion Criteria		Impaired Fasting Glucose (IFG) defined as 100 mg/dl (5.6mmol/l) ≥ FPG <125 mg/dl (6.9mmol/l) Impaired glucose tolerance (IGT) defined as 2-hour plasma glucose value 140 mg/dl - 199 mg/dl (7.8-11.1mmol/l) Prediabetes defined as IFG or IGT Diabetes defined as FPG ≥126 mg/dl (7mmol/l) or a 2-hour plasma glucose value ≥200 mg/dl (11.1mmol/l) -Timing of postnatal test: Performed at 6 weeks, 6 months and 1 year following delivery -Location of postnatal test (primary/secondary care): Unclear -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were the index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals from the study explained: Yes Other information NR: Not reported Only data for diabetes has been extracted as cut-offs for other outcomes do not match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	Women recruited from July 2007 to May 2009 who had a diagnosis of gestational diabetes Exclusion Criteria Women with arterial hypertension, renal disease, liver disease, thyroid disorders or other endocrine or chronic diseases				
Full citation Malinowska-Polubiec,A., Sienko,J., Lewandowski,Z., Czajkowski,K., Smolarczyk,R., Risk factors of abnormal carbohydrate metabolism after pregnancy complicated by gestational diabetes mellitus, Gynecological Endocrinology, 28,	Sample size Number with gestational diabetes: NR Number with postnatal test: 155 Characteristics Maternal age in years 19-48 Ethnicity White: 100%	Tests 2 hour 75g OGTT	Methods Gestational diabetes criteria: NR Outcomes: IFG, IGT, Diabetes Outcome definitions: WHO 1999 criteria. IFG defined as FPG ≥6.1 and <7.0mmol/l and normal 2 hour glucose level. IGT defined as 2 hour glucose ≥7.8 and <11.1mmol/l. Diabetes defined as FPG ≥7.0mmol/l or 2 hour glucose ≥11.1mmol/l	Results Incidence data IFG: 28/155 (18.1%) IGT: 31/155 (30%) Diabetes: 23/155 (14.8%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
360-364, 2012 Ref Id 177475 Country/ies where the study was carried out Poland Study type Case-control study Aim of the study To explore risk factors and to evaluate the risk of glucose intolerance and diabetes in women with a history of gestational diabetes Study dates All women delivered between 1998 and 2008 Source of funding NR	Parity, n (%) Multiparas: 26/155 (16.8%) Family history of diabetes (%) NR BMI NR Macrosomic infant delivered NR Medication during pregnancy NR * The characteristics above are of those who completed the postnatal test Inclusion Criteria - History of pregnancy complicated by gestational diabetes - At least the last pregnancy and		Timing of postnatal test: 6 months-10 years Location of postnatal test: Unclear Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	delivery managed in the Department of Obstetrics and Gynecology - The will to participate in the study - Signed informed consent Exclusion Criteria - Ongoing pregnancy at the onset of the study				
Full citation Rivas,A.M., Gonzalez,N., Gonzalez,J., High frequency of diabetes in early post-partum assessment of women with gestational diabetes mellitus, Diabetes and Metabolic Syndrome: Clinical Research	Sample size Number with gestational diabetes: 169 Number with postnatal test: 117 (69.2%) Characteristics Maternal age in years, mean (SD) 32.14 (6.76) Race/ethnicity	Tests 75g 2 hour OGTT	Methods -Gestational diabetes criteria: diagnosed using the Third International Gestational Diabetes Conference -Outcomes: IFG, IGT, diabetes -Outcome definitions: ADA 1997 criteria. Cut-offs not reported in article but extracted from a reference article. IFG defined as	Results Incidence data IFG: 14/117 (11.97%) Diabetes: 22/117 (18.80%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
and Reviews, 1, 159-165, 2007 Ref Id 179701 Country/ies where the study was carried out Venezuela Study type Prospective cohort study Aim of the study To determine the early glucose tolerance impairment, insulin resistance, and the association of other components of the metabolic syndrome in women with previous gestational diabetes Study dates All women were diagnosed with gestational diabetes and gave birth ('resolved pregnancy')	NR Parity, mean (SD) 3.4 (2.47) Family history of diabetes (%) 62.39 Prepregnancy BMI in kg/m2, mean (SD): 28.88 (4.97) Macrosomic infant delivered, % 23.93 Insulin use during index pregnancy (%) 36.75 * The characteristics above are of those who completed the postnatal test		fasting ≥6.1mmol/l and <7.0mmol/l. IGT defined as 2 hour glucose ≥7.8 and <11.1mmol/l. Diabetes defined as fasting ≥7mmol/l or 2 hour glucose ≥11.1mmol/l -Timing of postnatal test: 2-4 months postpartum -Location of postnatal test (primary/secondary care): Diabetes and Pregnancy Unit -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information Only data for diabetes and IFG has been

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
between September 1998 and September 2005 Source of funding Supported by research grant from Scientific and Humanistic Council of the University of Carabobo	Inclusion Criteria Patients referred to the University of Carabobo Diabetes and Pregnancy Unit diagnosed with gestational diabetes Exclusion Criteria NR				extracted as cut-off for IGT in article does not match the WHO criteria
Full citation Costa,A., Carmona,F., Martinez-Roman,S., Quinto,L., Levy,I., Conget,I., Post-partum reclassification of glucose tolerance in women previously diagnosed with gestational diabetes mellitus, Diabetic Medicine, 17, 595-598, 2000 Ref Id 180818	Sample size 120 women with previous gestational diabetes Characteristics Maternal age in years, mean (SD) In women with normal glucose tolerance: 33.9 (4.12) In women with abnormal glucose tolerance (IGT or diabetes): 36 (5.8)	Tests 2 hour 75g OGTT	Methods -Once breast feeding had finished, an OGTT was performed in 120 women with previous gestational diabetes. They were classified according to the WHO 1985 and ADA 1997 criteria (only ADA data extracted for this review) -Gestational diabetes criteria: 50g, 1 hour OGTT at the second trimester of gestation (22-26 weeks' gestation). A second test, at the third trimester (30-34 weeks' gestation) was performed when the former was normal. Women with a 1	Results Incidence data IFG: 4/120 (3%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No, exclusion criteria not reported 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard:

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Country/ies where the study was carried out Spain Study type Retrospective cohort study Aim of the study To evaluate postnatal screening based on FPG versus OGTT in Caucasian women with previous gestational diabetes Study dates All women delivered between 1997-1998 Source of funding Not reported	Race/ethnicity Caucasian (100%) Parity % NR Family history of diabetes (%) NR BMI (kg/m2), mean (SD): In women with normal glucose tolerance: 25.1 (4.3) In women with abnormal glucose tolerance (IGT or diabetes): 28.5 (6.3) Macrosomic infant delivered NR Insulin use during pregnancy (%) NR Inclusion Criteria -Caucasian women with a recent history of gestational		hour plasma glucose >7.8mmol/l underwent a 100g 3 hour antenatal OGTT and were classified as having gestational diabetes according to the Third International Workshop Conference on gestational diabetes recommendations -Outcomes: normal glucose tolerance, IFG, diabetes -Outcome definitions: Based on the FPG, the ADA 1997 criteria was used. Normal glucose tolerance <6.1mmol/l, IFG 6.1-6.9mmol/l and diabetes >7.0mmol/l. -Timing of postnatal test: 2-12 months after delivery -Location of postnatal test (primary/secondary care): Hospital -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		Yes, whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test? (that is, the index test did not form part of the reference standard): No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were the index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals from the study explained: Yes Other information Only data for IFG has been extracted as cut-off for diabetes does not exactly match the WHO criteria

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	diabetes, who gave written consent were studied after delivery during the period 1997-1998 Exclusion Criteria NR				
Full citation Aberg,A.E., Jonsson,E.K., Eskilsson,I., Landin-Olsson,M., Frid,A.H., Predictive factors of developing diabetes mellitus in women with gestational diabetes, Acta Obstetricia et Gynecologica Scandinavica, 81, 11-16, 2002 Ref Id 180886 Country/ies where the study was carried out Sweden	Sample size Number with gestational diabetes: 315 Number with postnatal test: 229 (73%) Characteristics Age in years, n -20: 1 20-24: 9 25-29: 79 30-34: 78 35-39: 48 40-44: 12 45-: 2 Ethnicity, n(%) NR Parity, n	Tests 75g 2-hour OGTT	Methods -Of 315 women with gestational diabetes, 229 underwent a further test at 1 year postpartum. The study compared maternal and fetal factors during pregnancy with the test value at follow-up. A control group of 153 women with a 2-hour test value below 7.8 mmol/l during pregnancy were invited to undergo a further test at 1 year postpartum and 60 (39%) accepted -Gestational diabetes criteria: The European Association for the Study of Diabetes (EASD) defining gestational diabetes as at least 9mmol/l as 2-hour values after a 75g	Results Incidence data Diabetes: 21/229 (9%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No, exclusion criteria not reported 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: NA (OGTT was performed and only 2-hour results were used). 5) Did the whole sample or a random selection of the sample receive verification using the reference standard:

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Study type Retrospective cohort study Aim of the study To investigate which factors in pregnancies complicated by gestational diabetes correlate with the risk of developing impaired glucose tolerance or diabetes at 1 year postpartum and to compare this risk in women with gestational diabetes and women with a normal oral glucose tolerance test during pregnancy Study dates All women with gestational diabetes delivered between 1991 and 1999	1: 75 2: 95 3: 41 4 : 18 Family history of diabetes NR BMI NR Macrosomic infant delivered NR Medication use NR * The characteristics above are of those who completed the postnatal test Inclusion Criteria - All gestational diabetes pregnancies delivered in Lund 1991-1999 Exclusion Criteria NR		OGTT -Outcomes: IGT, diabetes -Outcome definitions: The WHO definition of IGT as a 2-hour capillary blood concentration after a 75g OGTT between 7.8 and 11mmol/l and a value above 11mmol/l is considered to represent diabetes (it is not clear whether the 1985 or 1999 WHO criteria were used but 2-hour values are the same for both the 1985 and 1999 criteria in terms of IGT and diabetes) -Timing of postnatal test: 1 year postpartum -Location of postnatal test (primary/secondary care): Unclear -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		Yes, whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA (only 2-hour results used) 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: NA (only 2 hour results were used) 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NR: Not reported Only the data for diabetes was extracted. IGT cut-off in this article does not exactly match the WHO criteria as only the 2

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Source of funding NR					hour value was used to define IGT
Full citation Albareda,M., de,Leiva A., Corcoy,R., Reproducibility of diabetes mellitus diagnosis (WHO 1999 criteria) in women, Acta Diabetologica, 41, 14-17, 2004 Ref Id 181194 Country/ies where the study was carried out Study type To be decided Aim of the study Study dates Source of funding	Sample size Characteristics Inclusion Criteria Exclusion Criteria	Tests	Methods	Results	Limitations Other information This article reports identical incidence data to those reported in Albareda 2003 - please refer to the evidence table for Albareda 2003 for details
Full citation Kwak,S.H., Choi,S.H., Jung,H.S., Cho,Y.M., Lim,S., Cho,N.H.,	Sample size n=843 Characteristics N (%)	Tests 2-hour 75g OGTT	Methods All pregnant women received a 50-g 1-hour glucose challenge test with a positive cutoff value of7.2 mmol/L. Screen-positive women	Results Incidence Incidence of type 2 diabetes @ 2 months	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Kim,S.Y., Park,K.S., Jang,H.C., Clinical and genetic risk factors for type 2 diabetes at early or late post partum after gestational diabetes mellitus, Journal of Clinical Endocrinology and Metabolism, 98, E744-E752, 2013 Ref Id 247599 Country/ies where the study was carried out Korea Study type Prospective cohort study Aim of the study To investigate the clinical and genetic risk factors that are associated with type 2 diabetes early or late post partum after a pregnancy complicated by gestational	NGT/IGT = 738 (87.5) Type 2 DM = 105 (12.5) Age at pregnancy, years±SD NGT/IGT = 31.3±3.8 Type 2 DM = 32.1±4.0 P= 0.065 Pre-pregnancy BMI, kg/m2±SD NGT/IGT = 22.7±3.5 Type 2 DM = 24.2±3.8 P= <0.001 Pregnancy BMI at OGTT, kg/m2±SD NGT/IGT = 27.1±3.3 Type 2 DM = 28.3±3.6 P= <0.001 Weight gain during pregnancy, kg±SD NGT/IGT = 11.0±4.4		underwent a 100-g oral glucose tolerance test (OGTT) using the Third International Workshop-Conference diagnostic criteria. After delivery, women who had had gestational diabetes were scheduled for a 75g OGTT at 2 months post partum and annually thereafter. Subjects were categorized into normal glucose tolerance (NGT), impaired glucose tolerance (IGT), and type 2 diabetes groups according to the American Diabetes Association 2012 criteria. A total of 843 women who underwent the 75g OGTT at 2 months post partum were enrolled.	post partum = 105/843 = 12.5% Incidence of type 2 diabetes @ Median 49 months (IQR 30-82) post partum (women were negative at previous test) = 88/370 = 23.8%	representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: NA 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard:NA 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: NA 11) Were the reference standard results interpreted without knowledge of the results of the index test: NA 12) Were the same clinical data available when the test results were interpreted as

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
diabetes. Study dates Recruitment between January 1996 and February 2003 and follow up until December 2010 Source of funding Korea Healthcare Technology R&D Project Ministry of Health and Welfare	Type 2 DM = 9.9±4.8 P= 0.023 Gestational week at diagnosis, wk±SD NGT/IGT = 26.4±3.0 Type 2 DM = 25.2±5.3 P= 0.030 Parity, n±SD NGT/IGT = 0.48±0.64 Type 2 DM = 0.49±0.68 P= 0.913 Family history of DM, % NGT/IGT = 39.7 Type 2 DM = 47.6 P= 0.132 Inclusion Criteria Women with gestational diabetes attending Cheil General Hospital, Seoul, Korea.				would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	Participants were followed up at either at Cheil General Hospital or Seoul National University Bundang Hospital, Seongnam, Korea Exclusion Criteria Women who had diabetes before pregnancy or positive results for GAD antibodies were excluded.				
Full citation Katreddy,M.V., Pappachan,J.M., Taylor,S.E., Nevill,A.M., Indusekhar,R., Nayak,A.U., Hemoglobin A_{1c} in early postpartum screening of women with gestational diabetes, World Journal of Diabetes, 4, 76-81, 2013	Sample size n=203/408(49.8%) Characteristics Mean age = 29 ± 4.6 years Ethnic origin = 142 Caucasians (70%) and 61 Other racial groups (Asian: 60, Afro-Caribbean: 2,	Tests 75g 2 hour OGTT was performed after a minimum of 8 h overnight fast.	Methods All women who were diagnosed with GDM, managed by diet/lifestyle modifications and/or medical treatment, in the combined antenatal diabetes clinic between January 2010 and August 2012, were offered postpartum screening in the 6th week postpartum visit. These women were given counselling by the diabetic team, during their antenatal follow up, regarding the implications of GDM	Results Incidence At 6 weeks post partum IFG = 11/203 (5.4%) Type 2 diabetes = 7/203 (3.5%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: NA

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Ref Id 306166 Country/ies where the study was carried out England Study type Retrospective cohort study Aim of the study To explore the utility of HbA_{1c} in the early post-partum screening of women with gestational diabetes in a large university hospital in the United Kingdom. Study dates January 2010 and August 2012 Source of funding Not reported	others: 9) BMI = 30 ± 6.4 kg/m2 (Caucasians: 32 ± 5.1 kg/m2 and Asians 26 ± 4.2 kg/m2) Inclusion Criteria Women who were diagnosed with GDM, managed by diet/lifestyle modifications and/or medical treatment, in a combined antenatal diabetes clinic who had 6 week postnatal OGTT and HbA_{1c} results available. Exclusion Criteria There were no exclusion criteria		diagnosis and the need for screening in the post-partum period. Along with the OGTT, HbA_{1c} estimation was undertaken as a part of the post-partum screening test. Data of the test results from participants were collected and they were grouped into categories according to the values as normal, impaired glycaemia or diabetes. FBG values less than 6.1 mmol/L was taken as normal; FBG values between 6.1 mmol/L and 6.9 mmol/L as impaired fasting glucose (IFG); and FBG ≥ 7.0 mmol/L as diabetes. The OGTT results were classified by the WHO criteria: normal glucose tolerance (FBG < 6.0 mmo/L and/or 2-h PPBG < 7.8 mmol/L); impaired glucose tolerance (FBG ≥ 6.1 mmol/L and < 7.0 mmol/L, and/or 2-h PPBG between 7.8 and 11.0 mmol/L); and diabetes (FBG ≥ 7.0 mmol/L and/or 2-h PPBG ≥ 11.1 mmol/L).		5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard:NA 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: NA 11) Were the reference standard results interpreted without knowledge of the results of the index test: NA 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information
Full citation Joseph,F.,	Sample size n=147/258	Tests 75g 2	Methods Gestational diabetes criteria:	Results Incidence data	Limitations NICE guidelines manual 2009: Appendix

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Photiou,V., Verma,A., Goenka,N., Davies,J., Clement-Jones,M., Casson,I., Identifying women with persistent abnormal glucose metabolism following gestational diabetes mellitus: Changing times, changing populations and changing needs, British Journal of Diabetes and Vascular Disease, 13, 31-36, 2013 Ref Id 248036 Country/ies where the study was carried out England Study type Retrospective cohort study Aim of the study To identify the percentage of women with DM	women with gestational diabetes attending the joint diabetes and pregnancy clinics at the Countess of Chester Hospital and the University Hospital Aintree/Liverpool Women's Hospital joint clinic during the study period who had complete glucose testing and demographic data available Characteristics Age >35 = 63/147 (43%) BMI >30 = 68/147 (46%) Ethnicity = Caucasian 132 (90%), Asian 9 (6%), Afro-Caribbean 3 (2%), Southeast Asian 3 (2%) Gestations lasting beyond first	hour OGTT	FPG ≥ 5.6 and 2hG ≥ 7.8mmol/L Outcomes: IFG, IGT, Diabetes Outcome definitions:WHO 1999 criteria IFG: fasting plasma glucose ≥ 6.1 mmol/L (110 mg/dL) and <7 mmol/L (126 mg/dL). IGT: fasting plasma glucose (if available) <7.0 mmol/L (126 mg/dL) AND 2 hour post 75g glucose drink of ≥ 7.8 mmol/L (140 mg/dL) and <11.1 mmol/L (200 mg/dL). Diabetes: a fasting plasma glucose concentration ≥7 mmol/L (or 126 mg/dL) or ≥ 11.1mmol/L (200mg/dL) 2 hours post 75g glucose drink. Timing of postnatal test; 6 weeks postpartum Location of postnatal test: Unclear Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No	At 6wks post partum based on OGTT results Incidence IFG = 23/147 = 15.6% Incidence IGT = 21/147 = 14.2% Incidence DM = 8/147 = 5.4% Accuracy data FPG ≥ 6.0mmol/l for detecting diabetes@ 6wks post partum TP: 8 FP: 13 FN: 0 TN: 126 Sensitivity, % (95% CI): 94.4(58.9 - 100.0)** Specificity, % (95% CI): 90.4 (88.1 - 90.7)** LR+ (95% CI): 9.80 (4.94 - 10.77)** LR- (95% CI): 0.06 (0.000 - 0.47)** *Diagnostic accuracy measures and CIs calculated by NCC-WCH technical team based on data reported in the article **0.5 has been added to each cell (TP, FN, FP, TN) for diagnostic accuracy calculations to take into account the	G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No, exclusion criteria not reported 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes, whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
and impaired glucose tolerance (IGT) that would be missed using the National Institute for Health and Clinical Excellence (NICE) recommendation to use fasting plasma glucose (FPG) alone (and not an oral glucose tolerance test (OGTT)) six weeks after delivery to identify persistently abnormal glucose metabolism in women with gestational diabetes mellitus. Study dates January 2003 and July 2010 Source of funding No specific grant from any funding agency in the public, commercial, or not-for-profit	trimester = none 62 (42%), one 45 (31%), two 22 (15%), three 10 (7%), four to nine 8 (5%) Bad obstetric history = 24/147 (16%) Previous big baby (birthweight > 4.5kg) = 18/147 (12%) Previous GDM = 19/147 (13%) Number of previous pregnancies with GDM = one 129 (88%), one 14 (10%), two 4 (3%) Week GDM diagnosed = <30wks 80 (54 %), 30-32 wks 22 (15%), 32-34wks 21 (14%), 34-36 wks 10 (7%) and >36 wks 14 (10%) Treated with Insulin = 77/147 (52%) Inclusion Criteria All women			zeros	interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: NA 14) Were withdrawals explained: NA Other information

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
sectors	included in the analysis had an OGTT at or after 6 weeks post-partum Exclusion Criteria Not reported				
Full citation Chew,W.F., Rokiah,P., Chan,S.P., Chee,W.S., Lee,L.F., Chan,Y.M., Prevalence of glucose intolerance, and associated antenatal and historical risk factors among Malaysian women with a history of gestational diabetes mellitus.[Erratum appears in Singapore Med J. 2013 Jan;54(1):58], Singapore Medical Journal, 53, 814-820, 2012	Sample size n=342 NGT n = 172 Isolated IGT n = 42 Isolated IFG n = 46 Combined IGT/IFG n = 29 T2DM n = 53 Characteristics Age (yrs) NGT = 37.6 ± 5.3 Isolated IGT = 37.7 ± 5.0 Isolated IFG = 38.9 ± 5.6 Combined IGT/IFG = 39.7 ± 6.8 T2DM = 39.4 ± 4.5 Weight (kg) NGT = 61.6 ±	Tests 75g 2-hour oral glucose tolerance test	Methods A standard 75g 2-hour oral glucose tolerance test (75g 2-hour OGTT) was performed after participants had fasted overnight for at least 8–12 hours. Results of the 75-g 2-hour OGTT were evaluated according to the 2002 WHO criteria for T2DM (FPG ≥ 7.0 mmol/L and/or 2-hour PG ≥ 11.1 mmol/L), isolated IGT (FPG < 5.6 mmol/L and 2-hour PG ≥ 7.8 mmol/L to < 11.1 mmol/L),(18) and the 2006 American Diabetes Association criteria for isolated IFG (FPG ≥ 5.6 mmol/L to < 7.0 mmol/L).(21) Combined IGT/IFG was defined as FPG ≥ 5.6 mmol/L to < 7.0 mmol/L and 2-hour PG ≥ 7.8 mmol/L to < 11.1 mmol/L. Anthropometric measurements,	Results @ 1-5 years Incidence IGT = 27/170 = 15.9% Incidence T2DM = 15/170 = 8.8% @ 6-10 years (women were negative at previous test) Incidence IGT = 7/94 = 7.5% Incidence T2DM = 21/94 = 22.3% @ 11-15 years (women were negative at previous test) Incidence IGT = 8/78 = 10.3% Incidence T2DM = 17/78 = 21.8%	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: NA 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Ref Id 308920 Country/ies where the study was carried out Malaysia Study type Descriptive Aim of the study Cross-sectional study to determine the prevalence of prediabetes (isolated IGT, isolated IFG and combined IGT/IFG) and Type 2 diabetes, as well as the associated antenatal and historical risk factors among women with PGDM being treated at the University Malaya Medical Centre Study dates Not stated Source of funding Malaysian	11.7a Isolated IGT = 63.5 ± 11.7c Isolated IFG = 63.4 ± 11.1d Combined IGT/IFG = 66.5 ± 11.9 T2DM = 73.3 ± 12.5a,c,d Height (m) NGT = 1.55 ± 0.06 Isolated IGT = 1.55 ± 0.06 Isolated IFG = 1.55 ± 0.06 Combined IGT/IFG = 1.53 ± 0.07 T2DM = 1.56 ± 0.05 BMI (kg/m2) NGT = 25.69 ± 4.85a,b Isolated IGT = 26.59 ± 4.84c Isolated IFG = 26.22 ± 4.33d Combined IGT/IFG = 28.53 ± 5.07b T2DM = 30.26 ± 4.62a,c,d		demographic, clinical and socioeconomic data were obtained.		reference standard:NA 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: NA 11) Were the reference standard results interpreted without knowledge of the results of the index test: NA 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Government Intesified in Prioritiy Areas grant	aNGT vs. T2DM ($p < 0.05$). bNGT vs. combined IGT/IFG ($p < 0.05$). cIsolated IGT vs. T2DM ($p < 0.05$). dIsolated IFG vs. T2DM ($p < 0.05$). Inclusion Criteria Women with previous gestational diabetes between 20–50 years of age recruited from the hospital's database of women with gestational diabetes using a systematic random sampling method. The diagnosis of gestational diabetes was made based on the 1985 criteria of the World Health Organization (WHO). The				

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	duration from the index pregnancy with gestational diabetes ranged from three months to 15 years postpartum. Exclusion Criteria Women currently pregnant were excluded				
Full citation Gingras,V., Tchernof,A., Weisnagel,S.J., Robitaille,J., Use of glycated hemoglobin and waist circumference for diabetic screening in women with a history of gestational diabetes, Journal of Obstetrics and Gynaecology Canada: JOGC, 35, 810-815, 2013 Ref Id 306038 Country/ies where	Sample size n=178/215 (see exclusions below) Characteristics Age, years =36.4 ± 4.8 Time since latest pregnancy, years = 3.5 ± 1.9 Ethnicity (n = 165) = Non-Hispanic white 156 (94.6), Other 9 (5.4) Waist circumference, cm = 91.4 ± 14.6 BMI, kg/m2 = 27.8 ± 6.5	Tests 75g 2hour OGTT Type 2 diabetes = FPG ≥ 7.0mmol/L and/or a 2h-PG ≥ 11.1 mmol/L. Impaired fasting glycemia = FPG ≥ 5.6 mmol/L and < 7.0 mmol/L Impaired glucose tolerance	Methods Women were recruited using databanks from the Régie de l'assurance maladie du Québec, the provincial health plan registry. Height, BMI and waist circumference were measured and waist circumference ≥ 88 cm was used as the cut-off for risk stratification in analyses. A 2-hour 75g OGTT was performed in the morning after an overnight fast. Plasma glucose was measured enzymatically. A_{1C} was determined using the National Glycated Haemoglobin Standardization	Results Women were tested at a mean 3.5 ± 1.9 years after their most recent pregnancy. @ mean 3.5 ±1.9 years post pregnancy Incidence Type 2 diabetes = 32/182 (18%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: NA 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
the study was carried out Canada Study type Prospective cohort study Aim of the study To examine the adequacy of glycated hemoglobin (A_{1C}) and waist circumference (WC) measurements to detect impaired glucose metabolism among women with prior gestational diabetes Study dates Index pregnancy between 2003 and 2010 and women recruited between October 2009 and August 2011 Source of funding Canadian Institute for Health	Inclusion Criteria Women aged ≥ 18 years from the greater Quebec City area, with a diagnosis of gestational diabetes made between April 2003 and June 2010, who were not pregnant at the time of the study, and who did not have type 1 diabetes Exclusion Criteria Participants on medication for type 2 diabetes or dyslipidemia (n = 8), with previous bariatric surgery (n = 1), or with missing laboratory measurements from the OGTT (n = 21). Women who were tested less than 6 months after their most recent pregnancy (n = 7)	= 2h-PG ≥ 7.8 mmol/L and < 11.0 mmol/L Pre-diabetes was defined as impaired fasting glycemia or impaired glucose tolerance. "Any glucose intolerance" included pre-diabetes and type 2 diabetes. An HbA_{1C} level ≥ 5.7% was used as the cut-off for sensitivity and			index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard:NA 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: NA 11) Were the reference standard results interpreted without knowledge of the results of the index test: NA 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Research (CIHR) and Fonds de la recherche en sante du Quebec (FRSQ)	were excluded to avoid any bias due to glycemic control during pregnancy on A_{1C} measures.	specificity analyses			
Full citation Myers,J.E., Hasan,X., Maresh,M.J.A., Post-natal assessment of gestational diabetes: fasting glucose or full glucose tolerance test?, Diabetic MedicineDiabet.Med., n/a-n/a, 2014 Ref Id 319499 Country/ies where the study was carried out England Study type Retrospective cohort study Aim of the study To determine the performance of a fasting plasma glucose sample	Sample size n = 629 Characteristics Median age at birth of child (years) = 33 (Range 18-45) Median BMI at booking (kg/m2) = 29 (Range 17-50) Inclusion Criteria Women who were diagnosed with gestational diabetes (after screening criteria were applied) and who underwent a 6 week postpartum OGTT. Exclusion Criteria Women who did not have a 6	Tests 6 week postpartum 75g 2 hour OGTT Diabetes = FPG ≥ mmol/l and or a 2h result ≥ 11.1mmol/l Impaired fasting glycaemia = FPG 6.1 - 6.9 mmol/l Impaired glucose tolerance = 2 hr results 7.8 -11.0 mmol/l Normal	Methods All women with gestational diabetes were offered a 6 week postpartum 75g 2 hour OGTT	Results Incidence @ median 44 days (IQR 42-50) post partum Incidence Type 2 diabetes = 30/629 = 4.8% Diagnostic accuracy of FPG ≥ 5.6 threshold to predict Type 2 diabetes Sensitivity = 76 Specificity = 80 LR +ve = 3.8 LR -ve = 0.3 ≥ 6.1 to predict Type 2 diabetes Sensitivity = 90 (74.4-96.5) Specificity = 91 (88.8-93.3) LR +ve = 10.4 (7.8-13.8) LR -ve = 0.11 (0.03-	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: The whole sample 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard:No 8) Was the execution of the index test

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
compared with a full oral glucose tolerance test for the detection Study dates January 2003 to May 2013 Source of funding None	week postpartum OGTT or test results	glucose tolerance = FPG ≤ 6.0 mmol/l and 2 hour result ≤ 7.7		0.32) ≥ 7.0 to predict Type 2 diabetes Sensitivity = 76 (59.1 - 88.2) Specificity = 91 LR +ve = 8.4 LR -ve = 0.26 ≥ 5.6 to predict IGT Sensitivity = 77 Specificity = 84 LR +ve = 4.8 LR -ve = 0.27 ≥ 7.0 to predict IGT Sensitivity = 61 Specificity = 93 LR +ve = 8.7 LR -ve = 0.42	described in sufficient detail to permit its replication: Yes 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: Yes Other information
Full citation Agarwal,M.M., Punnose,J., Dhatt,G.S., Gestational diabetes: implications of variation in post-partum follow-up criteria, European Journal of Obstetrics,	Sample size Number with gestational diabetes: 1641 Number with postnatal test: 549 (33.5%) Characteristics Maternal age in years, mean (range)	Tests 2-hour 75g OGTT	Methods - During a 5-year period, 549 women underwent the 2-hour 75g OGTT. They were classified by the criteria of WHO (1985), the ADA (1997, fasting glucose) and the revised WHO (1999) - Gestational diabetes criteria: Women underwent an antenatal 100g 3-hour	Results Incidence data (by ADA) Normal glucose tolerance: 462/549 (84.2%) Impaired glucose tolerance: - Impaired fasting glucose: 51/549 (9.3%) Diabetes: 36/549 (6.6%) Incidence data (by WHO	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (exclusion criteria not reported) 3) Was the reference standard likely to classify the target condition correctly:

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Gynecology, and Reproductive Biology, 113, 149-153, 2004 Ref Id 179392 Country/ies where the study was carried out United Arab Emirates Study type Retrospective cohort study Aim of the study To compare the recommendations of the ADA with those of the WHO for evaluating women with gestational diabetes after birth Study dates All women underwent antenatal OGTT during a 5-year period (January 1998-December 2002) Source of funding	32 Ethnicity, n (%) Arabs: 78.8% Indian National: 20.5% Parity NR Family history of diabetes NR BMI NR Macrosomic infant delivered NR Medication during pregnancy NR Inclusion Criteria -Pregnant women attending routine obstetric clinics at the Al Ain Hospital, Al Ain, United Arab Emirates (UAE)		OGTT and diagnosis of gestational diabetes was made using the ADA criteria. Cut-offs were not reported in the article but extracted from a reference article - at least two glucose measurements ≥ the thresholds of fasting 5.3mmol/l, 1 hour 10.0mmol/l, and 2 hours 8.6mmol/l -Outcomes: Normal glucose tolerance, IGT, IFG, Diabetes -Outcome definitions: ADA 1997 criteria (based on FPG values only): normal fasting glucose FPG <6.1; impaired fasting glucose FPG 6.1-6.9mmol/l; and diabetes FPG>/=7mmol/l WHO 1999 criteria: normal glucose tolerance FPG <6.1mmol/l and 2-hour PG <7.8 mmol/l; IGT FPG <7mmol/l and 2-hour PG 7.8-11.0mmol/l; diabetes FPG>/=7mmol/l and/or 2-hour PG >/=11.1mmol/l; and IFG FPG 6.1-6.9mmol/l -Timing of postnatal test: 4-8 weeks after birth	1999) Normal glucose tolerance: 385/549 (70.1%) Impaired glucose tolerance: 84/549 (15.3%) Impaired fasting glucose: 30/549 (5.5%) Diabetes: 50/549 (9.1%) The difference for diabetes between the two criteria was not statistically significant (P=0.1) Accuracy data FPG>/=7.0mmol/l(126mg/dl) for detecting diabetes* TP: 36 FP: 0** FN: 14 TN: 499 Sensitivity, % (95% CI): 72(64.4-72.0) Specificity, % (95% CI): 100 (NC**) LR+ (95% CI): 72000*** LR- (95% CI): 0.280 (0.280-0.359) FPG>=6.1mmol/l for	Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
NR	Exclusion Criteria NR		-Location of postnatal test (primary/secondary care): Routine obstetric clinics at the Al Ain Hospital -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No	detecting diabetes* TP: 42 FP: 45 FN: 8 TN: 454 Sensitivity, % (95% CI): 84(71.7-92.1) Specificity, % (95% CI): 91 (89.7-91.8) LR+ (95% CI): 9.315 (6.995-11.230) LR- (95% CI): 0.176 (0.086-0.315) FPG <7mmol/l for detecting IGT* TP: 84**** FP: 429**** FN: 0**** TN: 36**** Sensitivity, % (95% CI): 99.4(94.2-100) Specificity, % (95% CI): 7.8 (6.9-7.9) LR+ (95% CI): 1.079 (1.012-1.086) LR- (95% CI): 0.075 (0-0.843) FPG <6.1mmol/l for detecting IGT*	Other information NC: Not calculable NR: Not reported Diagnostic accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html Reference article from which cut-offs for gestational diabetes (ADA criteria) were extracted: http://cdn.intechopen.com/pdfs/23174/InTech-Gestational_diabetes_evidence_based_screening_diagnosis_and_treatment.pdf

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				TP: 69 FP: 393 FN: 15 TN: 72 Sensitivity, % (95% CI): 82.1 (73.2-89.0) Specificity, % (95% CI): 15.5 (13.9-16.7) LR+ (95% CI): 0.972 (0.850-1.069) LR- (95% CI): 1.153 (0.656-1.929) FPG 6.1-6.9 for detecting IFG* TP: 30**** FP: 21**** FN: 0**** TN: 498**** Sensitivity, % (95% CI): 98.4(85.2-100) Specificity, % (95% CI): 95.9 (95.1-96) LR+ (95% CI): 23.796 (17.298-24.762) LR- (95% CI): 0.017 (0-0.156) *Diagnostic accuracy measures and CIs calculated by NCC-WCH technical team based on data reported	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				in the article **The specificity was fixed at 100% as all the 2 hour 75g OGTTs with negative results (FPG<7.0mmol/l and 2 hour plasma glucose <11.1mmol/l) will necessarily have an FPG <7.0mmol/l which means it is not possible to have a false positive ***Specificity was treated as 99.999% instead of 100% in order to calculate the LR **** 0.5 has been added to each cell (TP, FN, FP, TN) for diagnostic accuracy calculations to take into account the zeros	
Full citation Conway,D.L., Langer,O., Effects of new criteria for type 2 diabetes on the rate of postpartum glucose intolerance in women with gestational diabetes,	Sample size Number with gestational diabetes: 1017 Number with postnatal test: 179 (18%) Characteristics Maternal age (years) NR	Tests 2-hour 75g OGTT	Methods - Women identified as having gestational diabetes were instructed to undergo a 75g, 2-hour glucose tolerance test 4-6 weeks after delivery. The results were retrospectively categorised with both the 1979 NDDG criteria and those recommended by the ADA	Results Incidence data ADA 1997 (based on 2-hour OGTT) Diabetes: 14/179 (7.8%) Accuracy data FPG >=7mmol/l for detecting diabetes TP: 12	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (exclusion criteria not reported) 3) Was the reference standard likely to classify the target condition correctly:

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
American Journal of Obstetrics and Gynecology, 181, 610-614, 1999 Ref Id 178989 Country/ies where the study was carried out USA Study type Retrospective cohort study Aim of the study To determine the impact of the 1997 ADA diagnostic criteria for diabetes on the rate of postnatal glucose intolerance in women with gestational diabetes Study dates All gestational diabetes women delivered between 1 January 1995 and 30 June 1997 Source of funding	Race/ethnicity NR Parity % NR Family history of diabetes (%) NR BMI: NR Macrosomic infant delivered NR Insulin use during pregnancy (%) NR Inclusion Criteria -Women with gestational diabetes who were delivered at University Hospital in San Antonia betwen 1 January 1995 and 30 June 1997 and who subsequently underwent glucose tolerance		- Gestational diabetes criteria: NDDG 1979 criteria - 50g, 1-hour glucose challenge test, either at 24-28 weeks' gestation or on entry to antenatal care in the presence of risk factors for diabetes. Glucose challenge test values >/=130mg/dl(7.2mmol/l) were considered abnormal and prompted performance of a glucose tolerance test (GTT) -Outcomes: Normal, IGT, IFG, diabetes -Outcome definitions: ADA 1997 - Normal: FPG<110mg/dl (6.1mmol/l) and 2-hour PG <140mg/dl (7.8mmol/l), IGT: 2-hour PG >/=140mg/dl (7.8mmol/l) and <200mg/dl (11.1mmol/l), IFG: FPG >/=110mg/dl (6.1mmol/l) and <126mg/dl (7mmol/l), diabetes: FPG >/=126mg/dl (7mmol/l)* or 2-hour PG >/=200mg/dl(11.1mmol/l) *Diagnosis of diabetes based on FPG alone requires that this criterion be confirmed on	FN: 2 FP: NR TN: NR Sensitivity, % (95% CI): 85.71 (57.19 to 98.22)* *Calculated by NCC-WCH technical team based on data reported in the article	Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
NR	testing >/=4 weeks' after delivery Exclusion Criteria NR		a second occasion -Timing of postnatal test: 4-13 weeks' after delivery (mean 7 +/- 2 weeks) -Location of postnatal test (primary/secondary care): Unclear -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No		Other information NR: Not reported Diagnostic accuracy measures and CIs calculated using http://statpages.org/confint.html Only data for diabetes has been extracted as the cut-off matches the WHO 1999 criteria.
Full citation Ferrara,A., Peng,T., Kim,C., Trends in postpartum diabetes screening and subsequent diabetes and impaired fasting glucose among women with histories of gestational diabetes mellitus: A report from the Translating Research Into Action for Diabetes (TRIAD)	Sample size Number with gestational diabetes: 14448 (901 women had more than one pregnancy) Number with postnatal test: 5524 (38.2%) Characteristics Maternal Age in years. % <25 : 5.4 25–35 : 63.0 ≥36 : 31.6 Ethnicity, %	Tests 2-hour 75g OGTT	Methods - A cohort study of 14448 gestational diabetes pregnancies delivered between 1995 and 2006. Postnatal screening was defined as performance of either an FPG or OGTT at least 6 weeks after delivery and within 1 year of delivery - Gestational diabetes criteria: NDDG criteria- 50g 1-hour oral challenge test and if abnormal (>=7.8mmol/l) 3-hour 100g OGTT. Gestational diabetes was diagnosed if the woman had ≥ 2 glucose values at or exceeding the following	Results Incidence data RESULTS FOR 1995-2006 Total number of gestational diabetes pregnancies during study period: 14448 (13,547 women) Total number of pregnancies with postnatal test results: 5524 (38.2%) Using the FPG results only (either performed alone or as part of the OGTT) Diabetes: 191/5524	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: No, % with FPG and % with OGTT not

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Study, Diabetes Care, 32, 269-274, 2009 Ref Id 153194 Country/ies where the study was carried out USA Study type Retrospective cohort study Aim of the study To investigate a population of women with gestational diabetes, including trends in impaired fasting glucose (IFG) or diabetes detected with postpartum screening and the proportion of women with diabetes or prediabetes identified by the FPG screen versus the proportion of women with these abnormal glucose values identified	Non-Hispanic white : 28.0 African American : 3.2 Asian : 31.3 Hispanic : 27.1 Other : 5.6 Unknown : 4.8 Parity, % 0 : 40.4 1 : 32.8 ≥2 : 26.8 Family history of diabetes NR Obese, % 8.9 Macrosomic infant delivered, % 13.8 Diabetes medication during pregnancy, % Insulin : 15.2 Glyburide : 13.9 Inclusion Criteria - Women with diagnosis of		thresholds: fasting, 105 mg/dl (5.8mmol/l); 1 hour, 190 mg/dl (10.6mmol/l); 2 hours, 165 mg/dl (9.2mmol/l); and 3 hours, 145 mg/dl (8.1mmol/l) -Outcomes: IFG, IGT, prediabetes, diabetes -Outcome definitions: name of criteria not reported, cut-offs similar to ADA 2003 criteria Impaired Fasting Glucose - IFG: defined as FPG ≥100 mg/dl (5.6mmol/l) but <126 mg/dl Impaired glucose tolerance- defined as a 2-hour plasma glucose value ≥140 mg/dl (7.8mmol/l) Prediabetes - IFG or IGT Diabetes - defined as an FPG >/=126 mg/dl (7mmol/l) or a 2-hour plasma glucose value ≥200 mg/dl (11.1mmol/l) -Timing of postnatal test: Performed between 6 weeks' and 1 year following delivery -Location of postnatal test (primary/secondary care): Unclear	(3.5%) RESULTS FOR 1995-1997 Total number of gestational diabetes pregnancies screened postpartum for 1995 - 1997: 564 Using the FPG results only of the 75g OGTT Diabetes: 32/564 (5.7%) RESULTS FOR 2004-2006 Total number of gestational diabetes pregnancies screened postpartum: 2,381 Using the FPG results only Diabetes: 80 /2381 (3.4%) Accuracy data RESULTS FOR 2006 FPG >/=7.0mmol/l for detecting diabetes TP: 4 FP: NR FN: 12 TN: NR	reported but postpartum screening was defined as performance of either an FPG or OGTT -therefore assuming that not all subjects had OGTT 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test? (that is, the index test did not form part of the reference standard): No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were the index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals from the study explained: Yes Other information -NR:Not reported

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
by the 75-g oral glucose tolerance test (OGTT) Study dates All women delivered between 1 January 1995 and 31 December 2006 Source of funding Funds from the Translating Research Into Action for Diabetes (TRIAD) study (which in turn was supported by the Centers for Disease Control and Prevention and the National Institute of Health of Diabetes and Digestive and Kidney Diseases)	gestational diabetes from a health provider - Only women who met the NDDG criteria of gestational diabetes Exclusion Criteria - Clinical diagnosis of gestational diabetes not documented in notes		-Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No	Sensitivity, % (95% CI) : 25 (7.27-52.38)* *Calculated by NCC-WCH technical team based on data reported in the article	-Only the data for diabetes has been extracted as the cut-offs for all other outcomes in this article do not match the WHO 1999 criteria -Diagnostic accuracy measures and CIs calculated using http://statpages.org/confint.html
Full citation Holt,R.I., Goddard,J.R., Clarke,P., Coleman,M.A., A	Sample size Number with gestational diabetes: 152 Number with	Tests 2 hour 75g OGTT	Methods -Gestational diabetes criteria: WHO criteria using a cut-off value of fasting plasma glucose >=7.0mmol/l	Results Incidence data OGTT Diabetes: 3/122 (2.5%) IGT: 3/122 (2.5%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
postnatal fasting plasma glucose is useful in determining which women with gestational diabetes should undergo a postnatal oral glucose tolerance test.[see comment], Diabetic Medicine,Diabet.Med., 20, 594-598, 2003 Ref Id 182147 Country/ies where the study was carried out UK Study type Retrospective cohort study Aim of the study To identify whether fasting plasma glucose at 6 weeks after delivery can identify women with an abnormal OGTT and therefore	postnatal test: 122 (80.3%) Characteristics Maternal Age in years (range) 31.1 (18.7-38.9) Ethnicity, % Caucasian: 86% Asian: 14% Parity, % NR Family history of diabetes NR Macrosomic infant delivered, % NR Diabetes medication during pregnancy, % NR Inclusion Criteria Women with gestational diabetes diagnosed according to the		or a 2hour value of >=7.8mmol/l -Outcomes: IFG, IGT, prediabetes, diabetes -Outcome definitions: WHO 1999 criteria. Cut offs not reported in article but extracted from a reference article: Normal (fasting <6.1mmol/l, 2-hour <7.8mmol/l implied), IFG (fasting >=6.1 and <7.0mmol/l and 2-hour <7.8mmol/l if measured), IGT (fasting <7.0mmol/l and 2-hour >=7.8 and <11.1mmol/l), Diabetes (fasting >=7mmol/l or 2-hour >=11.1mmol/l) -Timing of postnatal test: 6 weeks' after delivery -Location of postnatal test (primary/secondary care): Princess Anne Hospital, Southampton -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No	IFG: 4/122 (3.3%) FPG Diabetes: 2/122 (1.6%) Accuracy data FPG >=7.0mmol/l for detecting diabetes* TP: 2** FN: 1 ** FP: 0 ** TN: 119 ** Sensitivity, % (95% CI): 62.5 (17.0-75.0) Specificity, % (95% CI): 99.6 (98.1-100.0) LR+ (95% CI): 150.000 (8.814-94371810.35) LR- (95% CI): 0.377 (0.250-0.846) FPG >=6.0mmol/l for detecting diabetes* TP: 3** FN: 0** FP: 7** TN: 112** Sensitivity, % (95% CI): 87.5 (31.5-100) Specificity, % (95% CI): 93.8 (91.9-94.2)	representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No, exclusion criteria not reported 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test? (that is, the index test did not form part of the reference standard): No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were the index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
determine which women should undergo a postnatal OGTT Study dates OGTTs performed between 1 May 2000 and 1 May 2002 Source of funding Not reported	WHO criteria Exclusion Criteria Not reported			LR+ (95% CI): 14 (3.884-17.143) LR- (95% CI): 0.133 (0-0.745) FPG <7mmol/l for detecting IGT* TP: 3** FN: 0** FP: 117** TN: 2** Sensitivity, % (95% CI): 87.5 (43.3-100) Specificity, % (95% CI): 2.1 (0.6-2.5) LR+ (95% CI): 0.894 (0.436-1.026) LR- (95% CI): 6 (0-92.847) FPG <6mmol/l for detecting IGT* TP: 0** FN: 3** FP: 112** TN: 7** Sensitivity, % (95% CI): 12.5 (0-68.5) Specificity, % (95% CI): 6.3 (5.8-8.1) LR+ (95% CI): 0.133 (0-0.745) LR- (95% CI): 14	when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals from the study explained: Yes Other information NR: Not reported Diagnostic accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				(3.884-17.143) FPG 6.0-6.9mmol/l for detecting IFG* TP: 4** FN: 0** FP: 4** TN: 114** Sensitivity, % (95% CI): 90 (40.5-100) Specificity, % (95% CI): 96.2 (94.1-96.6) LR+ (95% CI): 23.8 (6.899-29.750) LR- (95% CI): 0.104 (0-0.633) *Diagnostic accuracy measures and CIs calculated by NCC-WCH technical team based on data reported in the article **0.5 has been added to each cell (TP, FN, FP, TN) for diagnostic accuracy calculations to take into account the zeros	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation Hunt,K.J., Conway,D.L., Who returns for postpartum glucose screening following gestational diabetes mellitus?, American Journal of Obstetrics and Gynecology, 198, 404-406, 2008 Ref Id 154107 Country/ies where the study was carried out USA Study type Prospective cohort study Aim of the study To compare the characteristics of women who did and did not return for postnatal screening and to attempt to determine the prevalence and type of postnatal impaired	Sample size Number with gestational diabetes: 707 Number with postnatal test: 400 (57%) Characteristics Age in years, mean (95% CI) 29.6 (29.0, 30.2) Ethnicity, %(95% CI) Mexican American: 94 (91.7, 96.4) Parity NR Family history of diabetes, % (95% CI) 71.4 (66.9, 75.8) Prepregnancy BMI (kg/m2), mean (95% CI)	Tests 75g 2-hour OGTT: 288/400 (72%) FPG only: 112/400 (28%)	Methods - All women with gestational diabetes were instructed to undergo a postnatal OGTT 4-6 weeks' after delivery. Failure to undergo testing by the time of the routine postnatal examination triggered an additional contact by the case-manager nurse -Gestational diabetes criteria: The majority of women with gestational diabetes (96%) completed both a 50 g, 1-hour glucose challenge test and a 100 g, 3-hour OGTT. Cut-offs used to diagnose gestational diabetes not reported in article -Outcomes: Diabetes, IGT, IFG -Outcome definitions: Diabetes was defined as the presence of a fasting glucose level of 126 mg/dl (7mmol/l) or greater and/or a 2-hour postload glucose level of 200 mg/dl (11.1mmol/l) or greater. IGT was defined as a 2-hour glucose level of 140-199	Results Incidence data OGTT Diabetes: 13/288 (4.5%) FPG only Diabetes: 5/112 (4.5%) Accuracy data FPG>=7.0mmol/l (126mg/dl) to detect diabetes* TP: 4 FP: 0** FN: 9 TN: 275 Sensitivity,% (95% CI):30.8 (12.7-30.8) Specificity, % (95% CI): 100 (NC**) LR+ (95% CI): 30800*** LR- (95% CI): 0.692 (0.692-0.881) *Diagnostic accuracy measures and CIs calculated by NCC-WCH technical team based on data reported in the article **The specificity was	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (exclusion criteria not reported) 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: No (only 288 completed the OGTT) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
glucose regulation in a programme designed to increase postnatal testing for diabetes Study dates All women delivered between 29 March 2001 and 31 August 2003 Source of funding American Diabetes Association Research Award, National Institute of Diabetes and Digestive and Kidney Diseases Award, and the TDI (University Health System Community Health Initiatives)	29.1 (28.5, 29.7) Prior macrosomia, % (95% CI) 18.5 (14.7, 22.4) Medication use, % (95% CI) Gestational diabetes medication, any: 19 (15.6, 23.4) Glyburide only: 9.3 (6.4, 12.1) Insulin: 10.3 (7.3, 13.2) * The characteristics above are of those who completed the postnatal test Inclusion Criteria - Women with gestational diabetes who delivered at the University		mg/dl (7.8mmol/l-11.1mmol/l) and IFG as a fasting plasma glucose level of 100-125mg/dl (5.6mmol/l-6.9mmol/l) -Name of criteria not reported in article but cut-offs match ADA 2003 -Timing of postnatal test: 4-6 weeks' after delivery -Location of postnatal test (primary/secondary care): secondary (hospital visits and in-home glucose testing using an oral glucose load when hospital visits were not possible) -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No	fixed at 100%, as all the 2-hour 75g OGTTs with negative results (FPG<7.0mmol/l and 2-hour plasma glucose <11.1mmol/l) will necessarily have an FPG <7.0mmol/l which means it is not possible to have a false positive result ***Specificity was treated as 99.999% instead of 100% in order to calculate the LR	10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information -NC: Not calculable -NR: Not reported -Only data for diabetes has been extracted as the cut-offs for other outcomes in this article do not match the WHO 1999 criteria -Diagnostic accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	Hospital in San Antonio from 29 March 2001 to 31 August 2003 Exclusion Criteria NR				
Full citation Kitzmiller,J.L., ng-Kilduff,L., Taslimi,M.M., Gestational diabetes after delivery: Short-term management and long-term risks, Diabetes Care, 30, S225-S235, 2007 Ref Id 157625 Country/ies where the study was carried out USA Study type Retrospective cohort study Aim of the study To evaluate the yield of postnatal 2 hour	Sample size 527 women with gestational diabetes who completed postnatal test Characteristics Age, years (range) NR Ethnicity, n (%) Asian Indian: 77/527 (15) Far East Asian: 94/527 (18) Southeast Asian: 154/527 (29) Hispanic: 96/527 (18) Non-Hispanic white (Caucasian: european, russian or middle eastern origin): 106/527	Tests 75g 2-hour OGTT	Methods - Study evaluated the yield of postnatal 2-hour 75g GTTs performed in clinical laboratories in a multi-ethnic population of women with gestational diabetes treated during 2000-2003 -Gestational diabetes criteria: diagnosed by private clinicians based on a 50g 1-hour glucose screening test value >199mg/dl (>11.1mmol/l) or a 100g 3-hour GTT with any two values >=95mg/dl fasting, 1 hour 180mg/dl, 2 hours 155 mg/dl and 3 hours 140mg/dl (5.3, 10.0, 8.6, 7.8 mmol/l, respectively) -criteria unamed in article but matches ADA -Outcomes: IFG, IGT, type 2 diabetes -Outcome definitions: Article	Results Incidence data, n (%) Diabetes: 25 (4.7) Accuracy data FPG>=7.0mmol/l (126mmol/l) to detect diabetes* TP: 4 FP: 0** FN: 21 TN: 502 Sensitivity, % (95% CI): 16 (6.5-16) Specificity, % (95% CI): 100 (NC**) LR+ (95% CI): 16000*** LR- (95% CI): 0.840 (0.840-0.940) *Diagnostic accuracy measures and CIs calculated by NCC-WCH technical team based on data reported	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (inclusion and exclusion criteria not reported) 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
75g glucose tolerance tests performed in clinical laboratories in a multi-ethnic population of women with gestational diabetes treated during 2000-2003 Study dates All women with gestational diabetes who were treated during 2000-2003 Source of funding NR	(20) Parity NR Family history of diabetes NR Prepregnancy BMI, %, kg/m2 BMI <25: 60.1 BMI 25-29.9: 25.6 BMI>/=30: 14.3 Macrosomic infant delivered NR Medication during pregnancy, % Medical nutrition therapy: 192/527 (36) Glyburide: 77/527 (15) Glyburide>insulin: 64/527 (12)** Insulin: 194/527 (37) * The characteristics above are of those who completed the postnatal test		states ADA 2003 criteria were used. Cut-offs not explicitly stated in article and have been extracted from the report of a WHO/IDF consultation. Normal: FPG<100mg/dl (5.6mmol/l) and 2 hour plasma glucose (PG) <140mg/dl(7.8mmol/l), IFG: FPG 100-125mg/dl (5.6-6.9mmol/l), IGT: 2-hour PG 140-199mg/dl (7.8-11.1mmol/l) and diabetes: FPG>=126mg/dl (7mmol/l) or 2-hour PG >=200mg/dl (11.1mmol/l) -Timing of postpnatal test: 6-21 weeks (timing depending on continuation of health insurance coverage) -Location of postnatal test (primary/secondary care): Clinical laboratories -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No	in the article **The specificity was fixed at 100%, as all the 2-hour 75g OGTTs with negative results (FPG<7.0mmol/l and 2-hour plasma glucose <11.1mmol/l) will necessarily have an FPG <7.0mmol/l which means it is not possible to have a false positive result ***Specificity was treated as 99.999% instead of 100% in order to calculate the LR	index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NC: Not calculable NR: Not reported IDF: International Diabetes Federation Data for diabetes only have been extracted as the cut-offs for other outcomes in the article do not match the WHO 1999 criteria Diagnostic accuracy measures and CIs

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	**Paper does not explain the use of the > sign but assuming this means glyburide followed by insulin Inclusion Criteria NR Exclusion Criteria NR				calculated using http://statpages.org/ctab2x2.html
Full citation Kousta,E., Lawrence,N.J., Penny,A., Millauer,B.A., Robinson,S., Dornhorst,A., de,Swiet M., Steer,P.J., Grenfell,A., Mather,H.M., Johnston,D.G., McCarthy,M.I., Implications of new diagnostic criteria for abnormal glucose homeostasis in women with previous	Sample size Number with gestational diabetes: 192 Number with postnatal test: 165 (85.9%) (27 of the 192 were excluded on the basis of having type 2 diabetes diagnosed after the index pregnancy) Characteristics Age in years, mean (SD) 36.6 (5.4)	Tests 75g 2 hour OGTT	Methods -Gestational diabetes criteria: At St Mary's gestational diabetes was diagnosed when the area under the plasma glucose curve exceeded 43 mmol/l/h during a 3 hour 75g OGTT. Elsewhere, diagnosis was based on the 2 hour plasma glucose, with all women exceeding WHO criteria for glucose intolerance during pregnancy of 7.8 mmol/l (although some centres adopted higher thresholds for clinical intervention). Most centeres used a modified O'Sullivan protocol	Results Incidence data FPG only IFG: 18/165 (10.9%) Diabetes: 19/165 (11.5%) OGTT IGT: 49/165 (29.7%) IFG: 7/165 (4.2%) Diabetes: 25/165 (15.2%) Accuracy data FPG>=7.0mmol/l for detecting diabetes TP: 19**	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: Yes 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard:

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
gestational diabetes, Diabetes Care, 22, 933-937, 1999 Ref Id 153415 Country/ies where the study was carried out UK Study type Retrospective cohort study Aim of the study To determine consequences of applying revised ADA 1997 and the WHO 1999 recommendations for the classification of glucose intolerance in women with previous gestational diabetes Study dates July 1997-June 1998 Source of funding	Ethnicity, n (%) European: 68 (35) South Asian (from India, Pakistan, Sri Lanka or Bangladesh): 56 (29) Afro-Caribbean: 32 (17) Other/mixed origin: 36 (19) Median Parity (range) 2 (1-8) Family history of diabetes, % NR BMI, kg/m2 28.1 +/-6.2 Macrosomic infant delivered NR Medication use during pregnancy, % insulin NR Inclusion Criteria -Women with		as a preliminary screen -Outcomes: Diabetes, IGT, IFG, Normal glucose tolerance -Outcome definitions: For ADA 1997, only FPG was used. Normal was defined as fasting <6.1, IFG was defined as FPG 6.1-6.9 and diabetes was defined as FPG >=7.0mmol/l. For WHO 1999, normal was defined as FPG <6.1mmol/l and 2 hour plasma glucose <7.8mmol/l, IFG was defined as FPG 6.1-6.9 and 2 hour plasma glucose <7.8mmol/l, IGT was defined as FPG <7.0mmol/l and 2 hour 7.8-11.0mmol/l and diabetes was defined as FPG >=7.0mmol/l or 2 hour plasma glucose >=11.1mmol/l -Timing of postnatal test: 1-86 months -Location of postnatal test (primary/secondary care): Unclear -Did study document a return to euglycaemia in the immediate days following	FN: 6** FP: 0** TN: 140** Sensitivity, % (95% CI): 75.0 (61.4-76.9) Specificity, % (95% CI): 99.6 (97.1-100) LR+ (95% CI): 211.500 (21.469-113730160.5) LR- (95% CI): 0.251 (0.231-0.397) *Diagnostic accuracy measures and CIs calculated by NCC-WCH technical team based on data reported in the article **0.5 has been added to each cell (TP, FN, FP, TN) for diagnostic accuracy calculations to take into account the zeros	Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information -Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Funded by a project grant from the UK Medical Research Council and through support from the Joint Research Standing Committee at St Mary's Hospital, London	previous gestational diabetes recruited retrospectively from 5 London Hospitals (St Mary's, Hammersmith and Queen Charlotte's, Chelsea and Westminster, Ealing, and Central Middlesex) Exclusion Criteria -Women with type 2 diabetes diagnosed since the index gestational diabetes pregnancy		delivery and before discharge: No		-NR: Not reported
Full citation McClean,S., Farrar,D., Kelly,C.A., Tuffnell,D.J., Whitelaw,D.C., The importance of postpartum glucose tolerance testing after	Sample size Number with gestational diabetes: 1189 Number with postnatal test: 985 (82.8%); 93 wom en experienced	Tests 75g 2-hour OGTT	Methods -Retrospective study of 985 pregnancies over a 10-year period in a mixed ethnic cohort of women who underwent follow-up glucose tolerance testing at 6 weeks' postpartum. Diagnosis obtained by OGTT was tested against that from the	Results Incidence data Normal: 713/985 (72%) Diabetes: 109/985 (11%) IGT: 114/985 (12%) IFG: 101/985 (10%) IGT and IFG: 52/985	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (exclusion criteria not reported)

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
pregnancies complicated by gestational diabetes, Diabetic Medicine, 27, 650-654, 2010 Ref Id 144569 Country/ies where the study was carried out UK Study type Retrospective cohort study Aim of the study To review postnatal glucose tolerance in women with gestational diabetes and evaluate the role of a formal 75g oral glucose tolerance test (OGTT) versus fasting plasma glucose (FPG) in screening for persistent abnormalities Study dates	gestational diabetes in two or more pregnancies during the study period Characteristics Maternal age at delivery, years (range) South Asian (Pakistani, Bangladeshi or Indian): 31 (27-35) White European: 32 (28-36) Whole group: 31 (27-35) Ethnicity, n(%) South Asian (Pakistani, Bangladeshi or Indian): 690/985 (71%) White European: 260/985 (26%) NR: 35/985 (4%)		fasting plasma glucose value -Gestational diabetes criteria: 75g OGTT at 24-28 weeks' gestation. Women were defined as having gestational diabetes if they fulfilled the WHO 1999 criteria for impaired fasting glucose (fasting plasma glucose >/=6.1mmol/l) and/or impaired glucose tolerance (2-hour post-challenge plasma glucose >=7.8mmol/l) -Outcomes: Normal, IFG, Diabetes -Outcome definitions: WHO 1999 cut-offs not reported in article but extracted from a reference article: Normal (fasting <6.1mmol/l, 2-hour <7.8mmol/l implied), IFG (fasting >=6.1 and <7.0mmol/l and 2-hour <7.8mmol/l if measured), IGT (fasting <7.0mmol/l and 2-hour >=7.8 and <11.1mmol/l), Diabetes (fasting >=7mmol/l or 2-hour >=11.1mmol/l) -Timing of postnatal test: 6 weeks after delivery	(5%) Accuracy data FPG >=7.0mmol/l for detecting diabetes* TP: 84** FN: 25** FP: 0** TN: 876** Sensitivity, % (95% CI): 76.8 (72.8-77.3) Specificity, % (95% CI): 99.9 (99.4-100) LR+ (95% CI): 1347.391 (129.872-710600901.4) LR- (95% CI): 0.232 (0.227-0.273) FPG >=6.1mmol/l for detecting diabetes* TP: 98 FN: 11 FP: 101 TN: 775 Sensitivity, % (95% CI): 89.9 (82.9-94.5) Specificity, % (95% CI): 88.5 (87.6-89.0) LR+ (95% CI): 7.798 (6.683-8.625)	3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
All women diagnosed with gestational diabetes between 1999 and 2008 Source of funding NR	Parity NR Family history of diabetes NR BMI NR Macrosomic infant delivered NR Medication during pregnancy NR * The characteristics above are of those who completed the postnatal test Inclusion Criteria - Women were included regardless of the number of		-Location of postnatal test (primary/secondary care): secondary care (antenatal care was in a hospital - assuming that participants returned for follow-up postnatal test at the same location) -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No	LR- (95% CI): 0.114 (0.062-0.195) FPG >=5.6mmol/l for detecting diabetes* TP: 106 FN: 3 FP: 222 TN: 654 Sensitivity, % (95% CI): 97.2 (91.7-99.3) Specificity, % (95% CI): 74.7 (74.0-74.9) LR+ (95% CI): 3.837 (3.525-3.957) LR- (95% CI): 0.037 (0.010-0.112) FPG >=5.1mmol/l for detecting diabetes* TP: 108 FN: 1 FP: 445 TN: 431 Sensitivity, % (95% CI): 99.1 (94.3-100) Specificity, % (95% CI): 49.2 (48.6-49.3) LR+ (95% CI): 1.950 (1.836-1.972) LR- (95% CI): 0.019 (0.001-0.116)	or intermediate test results reported: Yes 14) Were withdrawals explained: Yes Other information -Diagnostic test accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html -NR: Not reported

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	pregnancies in which they fulfilled the defined criteria for gestational diabetes - Women with incomplete data for antenatal items examined were included if complete postnatal OGTT results were available Exclusion Criteria - NR			FPG<7mmol/l for detecting IGT* TP: 114** FN: 0** FP: 787** TN: 84** Sensitivity, % (95% CI): 99.6 (95.4-100) Specificity, % (95% CI): 9.7 (9.1-9.7) LR+ (95% CI): 1.102 (1.051-1.108) LR- (95% CI): 0.045 (0.000-0.498) FPG <=6mmol/l for detecting IGT* TP: 62 FN: 52 FP: 724 TN: 147 Sensitivity, % (95% CI): 54.4 (45.8-62.9) Specificity, % (95% CI): 16.9 (15.7-18.0) LR+ (95% CI): 0.654 (0.543-0.767) LR- (95% CI): 2.703 (2.062-3.445) FPG <=5.5mmol/l for	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				detecting IGT* TP: 36 FN: 78 FP: 621 TN: 250 Sensitivity, % (95% CI): 31.6 (23.8-40.4) Specificity, % (95% CI): 28.7 (27.7-29.9) LR+ (95% CI): 0.443 (0.328-0.576) LR- (95% CI): 2.384 (1.995-2.755) FPG <=5.0mmol/l for detecting IGT* TP: 17 FN: 97 FP: 415 TN: 456 Sensitivity, % (95% CI): 14.9 (9.3-22.7) Specificity, % (95% CI): 52.4 (51.6-53.4) LR+ (95% CI): 0.313 (0.191-0.487) LR- (95% CI): 1.625 (1.448-1.758) FPG 6.1-6.9mmol/l for detecting IFG*	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				TP: 49** FN: 0** FP: 66** TN: 870** Sensitivity, % (95% CI): 99 (89.9-100) Specificity, % (95% CI): 92.9 (92.4-93) LR+ (95% CI): 13.949 (11.863-14.197) LR- (95% CI): 0.011 (0-0.109) FPG 5.6-6.9mmol/l for detecting IFG* TP: 49** FN: 0** FP: 195** TN: 741** Sensitivity, % (95% CI): 99 (89.7-100) Specificity, % (95% CI): 79.1 (78.6-79.2) LR+ (95% CI): 4.745 (4.199-4.805) LR- (95% CI): 0.013 (0-0.131) FPG 5.1-6.9mmol/l for detecting IFG* TP: 49** FN: 0**	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				FP: 420** TN: 516** Sensitivity, % (95% CI): 99 (89.6-100) Specificity, % (95% CI): 55.1 (54.6-55.2) LR+ (95% CI): 2.206 (1.975-2.231) LR- (95% CI): 0.018 (0-0.190) FPG <=5mmol/l to 6.9 for detecting IFG* TP: 49** FN: 0** FP: 852** TN: 84** Sensitivity, % (95% CI): 99 (89.9-100) Specificity, % (95% CI): 9 (8.5-9.1) LR+ (95% CI): 1.088 (0.983-1.100) LR- (95% CI): 0.111 (0-1.185) *Diagnostic accuracy measures and CIs calculated by NCC-WCH technical team based on data reported in the article	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				**0.5 has been added to each cell (TP, FN, FP, TN) for diagnostic accuracy calculations to take into account the zeros	
Full citation Megia,A., Naf,S., Herranz,L., Serrat,N., Yanez,R.E., Simon,I., Vendrell,J., The usefulness of HbA_{1c} in postpartum reclassification of gestational diabetes, BJOG: An International Journal of Obstetrics and Gynaecology, 119, 891-894, 2012 Ref Id 181892 Country/ies where the study was carried out Spain Study type Prospective cohort	Sample size Number with postnatal test: 364 with gestational diabetes attending postnatal assess ment Characteristics Age in years (range) For women with diabetes: 36 (30.5-39.75) For women without diabetes: 3 (2-5) Ethnicity (%) European: 91.5% Arabic: 5.5% Hispanic: 1.6% Others: 1.4% Parity	Tests 75g 2 hour OGTT, HbA_{1c}	Methods -Gestational diabetes criteria: NDDG criteria -Outcomes: Normal, IFG, IGT, Diabetes -Outcome definitions: WHO 1999 cut-offs not reported in article but extracted from a reference article: Normal (fasting <6.1mmol/l, 2-hour <7.8mmol/l implied), IFG (fasting >=6.1 and <7.0mmol/l and 2-hour <7.8mmol/l if measured), IGT (fasting <7.0mmol/l and 2-hour >=7.8 and <11.1mmol/l), Diabetes (fasting >=7mmol/l or 2-hour >=11.1mmol/l) -Timing of postnatal test: within the first year postpartum -Location of postnatal test (primary/secondary care):	Results Incidence data OGTT IFG/IGT or both: 89/364 (24.5%) Diabetes: 12/364 (3.3%) FPG Diabetes: 7/364 (1.9%) HbA_{1c} of 6.5% or more (diabetes): 2/364 (0.5%) Accuracy data FPG >/=7.0mmol/l for detecting diabetes TP: 7 FN: 5 FP: NR TN: NR Sensitivity, % (95% CI): 58.33 (27.67-84.83)*	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No, exclusion criteria not reported 3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test? (that is, the index test did not form part of the reference standard): No

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
study Aim of the study To analyse whether the use of HbA_{1C} may be useful in the postpartum reclassification of women with gestational diabetes in a large cohort of women Study dates Women returned for post-delivery study visit between January 2006 and March 2011 Source of funding Supported by grants from the Instituto de Salud Carlos III	NR Family history of diabetes NR BMI in kg/m2 Pre-gravid BMI For women with diabetes: 30.1 (26.8-32.7) For women without diabetes: 24.8 (22.2-25.6) Postpartum BMI For women with diabetes: 29.2 (26.4-33.5) For women without diabetes: 25.7 (22.7-30.2) Macrosomic infant delivered NR Medication during pregnancy, % insulin For women with diabetes: 100% For women without diabetes: 47%		Not reported -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No	Sensitivity and specificity of HbA_{1C} at various cut-off levels according to the OGTT criteria HbA_{1C} 5.3% Sensitivity (%): 91.67** Specificity (%): 72.44** LR+ : 3.33*** LR-: 0.11*** HbA_{1C} 5.4% Sensitivity (%): 75.00** Specificity (%): 82.67** LR+ : 4.33*** LR-: 0.30*** HbA_{1C} 5.5% Sensitivity (%): 66.67** Specificity (%): 88.07** LR+ : 5.59*** LR-: 0.38*** HbA_{1C} 5.6% Sensitivity (%): 41.67** Specificity (%): 92.05** LR+ : 5.24*** LR-: 0.63*** HbA_{1C} 5.7% Sensitivity (%): 41.67** Specificity (%): 96.31** LR+ : 11.29*** LR-: 0.61***	8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were the index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: Yes 13) Were uninterpretable, indeterminate or intermediate test results reported: Yes 14) Were withdrawals from the study explained: Yes Other information -Diagnostic test accuracy measures and CIs calculated using http://statpages.org/confint.html -NR: Not reported

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
	Inclusion Criteria Women that returned for the post-delivery study visit within the first year postpartum between January 2006 and March 2011 and had HbA_{1C} measured at the time of the postnatal 2 hour 75g OGTT Exclusion Criteria Not reported			HbA_{1C} 5.8% Sensitivity (%): 41.67** Specificity (%): 98.86** LR+ : 36.55*** LR-: 0.59*** HbA_{1C} 5.9% Sensitivity (%): 33.33** Specificity (%): 100** LR+ : 33330**** LR-: 0.67*** HbA_{1C} 6.0% Sensitivity (%): 25.00** Specificity (%): 100** LR+ : 25000**** LR-: 0.75*** HbA_{1C} 6.5% Sensitivity (%): 16.67** Specificity (%): 100** LR+ : 16670**** LR-: 0.83*** HbA_{1C} >=5.7% to diagnose any kind of glucose intolerance**: 13.5% and 97.3% respectively Sensitivity (%): 13.5** Specificity (%): 97.3** LR+ : 5*** LR-: 0.89*** Area under the ROC	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				curve For diagnosis of diabetes: 0.870 For diagnosis of any kind of glucose intolerance: 0.674 *Diagnostic accuracy measure and CI calculated by NCC-WCH technical team based on data reported in the article **Confidence intervals not reported ***LRs calculated by the NCC-WCH technical team based on data reported in the article. CIs non-calculable. ****Specificity was treated as 99.999% instead of 100% in order to calculate the LR .	
Full citation Jacob Reichelt,A.A., Ferraz,T.M., Rocha Oppermann,M.L., Costa e Forti, Duncan,B.B., Fleck,Pessoa E., Schmidt,M.I.,	Sample size Number with gestational diabetes: 159 Number with postnatal test: 117 (73.6%) Characteristics Age in years,	Tests 2 hour 75g OGTT	Methods -Gestational diabetes criteria: Not reported -Outcomes: Diabetes, impaired fasting glucose (IFG), impaired glucose tolerance (IGT) -Outcome definitions: Name	Results Incidence data FPG only Diabetes: 8/117 (6.8%) OGTT Diabetes: 9/117 (7.7%)	Limitations NICE guidelines manual 2009: Appendix G: the QUADAS tool for studies of diagnostic test accuracy 1) Was the spectrum of participants representative of the patients who will receive the test in practice: Yes 2) Were selection criteria clearly described: No (exclusion criteria not reported)

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Detecting glucose intolerance after gestational diabetes: inadequacy of fasting glucose alone and risk associated with gestational diabetes and second trimester waist-hip ratio, Diabetologia, 45, 455-457, 2002 Ref Id 183753 Country/ies where the study was carried out Brazil Study type Prospective cohort study Aim of the study To evaluate glucose alterations and associated risk factors 4-8 years after pregnancy in a subsample of the Brazilian Study of Gestational Diabetes	mean (range) NR Ethnicity, n(%) NR Parity NR Family history of diabetes NR BMI NR Macrosomic infant delivered NR Medication use NR Inclusion Criteria - All women with gestational diabetes and a randomly assigned sample of control subjects from a large cohort in Brazil (case-cohort study design assumed but not clearly reported in		of criteria not explicitly reported but assumed to be WHO based on cut-offs reported in article. The following cut-off levels were reported in the article: diabetes was defined as FPG >/=7.0mmol/l or 2 hour >/=11.1mmol/l, IGT was defined as FPG <7.0mmol/l and 2 hour >/=7.8 and IFG was defined as FPG >/=6.1mmol/l and 2 hour <7.8mmol/l. -Timing of postnatal test: 4-8 years after index pregnancy -Location of postnatal test (primary/secondary care): NR -Did study document a return to euglycaemia in the immediate days following delivery and before discharge: No	Accuracy data FPG>=7mmol/l for detecting type 2 diabetes* TP: 8 FP: 0** FN: 1 TN: 108 Sensitivity,% (95% CI): 88.9 (59.8-88.9) Specificity, % (95% CI): 100 (NC**) LR+ (95% CI): 88900*** LR- (95% CI): 0.111(0.111-0.412) FPG>=6.1mmol/l for detecting type 2 diabetes* TP: 8 FP: 12 FN: 1 TN: 96 Sensitivity,% (95% CI): 88.9 (53.2-99.4) Specificity, % (95% CI): 88.9 (85.9-89.8) LR+ (95% CI): 8 (3.778-9.714) LR- (95% CI): 0.125(0.007-0.545) FPG <7.0mmol/l for	3) Was the reference standard likely to classify the target condition correctly: Yes 4) Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests: Yes 5) Did the whole sample or a random selection of the sample receive verification using the reference standard: Yes (whole sample) 6) Did participants receive the same reference standard regardless of the index test result: Yes 7) Was the reference standard independent of the index test i.e. the index test did not form part of the reference standard: No 8) Was the execution of the index test described in sufficient detail to permit its replication: NA 9) Was the execution of the reference standard described in sufficient detail to permit its replication: Yes 10) Were index test results interpreted without knowledge of the results of the reference standard: Unclear 11) Were the reference standard results interpreted without knowledge of the results of the index test: Unclear 12) Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice: No 13) Were uninterpretable, indeterminate

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
Study dates Not reported Source of funding Foundation for the Support of Research of the State of Rio Grande do Sul, Fund for the Support of Research of the Hospital de Clinicas de Porto Alegre, a Centers of Excellence Grant, and Bristol-Myers Squibb Foundation	article) Exclusion Criteria NR			detecting IGT* TP: 39**** FP: 70**** FN: 0**** TN: 8**** Sensitivity,% (95% CI): 98.8 (90.6-100) Specificity, % (95% CI): 10.8 (6.6-11.4) LR+ (95% CI): 1.107 (0.970-1.129) LR- (95% CI): 0.116 (0.000-1.430) FPG <6.1mmol/l for detecting IGT* TP: 30 FP: 67 FN: 9 TN: 11 Sensitivity,% (95% CI): 76.9 (66.1-87.2) Specificity, % (95% CI): 14.1 (8.7-19.2) LR+ (95% CI): 0.896 (0.724-1.080) LR- (95% CI): 1.636 (0.665-3.908) FPG >=6.1mmol/l to 6.9mmol/l for detecting IFG* TP: 3**** FP: 9****	or intermediate test results reported: Yes 14) Were withdrawals explained: NA Other information NC: Not calculable NR: Not reported Only data for diabetes has been extracted as the cut-offs for other outcomes do not exactly match the WHO 1999 criteria Diagnostic accuracy measures and CIs calculated using http://statpages.org/ctab2x2.html

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				FN: 0**** TN: 105**** Sensitivity,% (95% CI): 87.5 (31.3-100) Specificity, % (95% CI): 91.7 (89.8-92.2) LR+ (95% CI): 10.592 (3.064-12.778) LR- (95% CI): 0.136 (0-0.765) *Diagnostic accuracy measures and CIs calculated by NCC-WCH technical team based on data reported in the article **The specificity was fixed at 100%, as all the 2-hour 75g OGTTs with negative results (FPG<7.0mmol/l and 2-hour plasma glucose <11.1mmol/l) will necessarily have an FPG <7.0mmol/l which means it is not possible to have a false positive ***Specificity was treated as 99.999% instead of 100% in order to calculate the LR ****0.5 has been added	

Bibliographic details	Participants	Tests	Methods	Outcomes and results	Comments
				to each cell (TP, FN, FP, TN) for diagnostic accuracy calculations to take into account the zeros	