Evidence Report/Technology Assessment

Number 210

Screening and Diagnosing Gestational Diabetes Mellitus

AHRQ
Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • **www.ahrq.gov**

Evidence-Based Practice

Evidence Report/Technology Assessment

Number 210

Screening and Diagnosing Gestational Diabetes Mellitus

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-2007-10021-I

Prepared by:
University of Alberta Evidence-based Practice Center
Edmonton, Alberta, Canada

Investigators:
Research Team:
Lisa Hartling, Ph.D.
Donna M. Dryden, Ph.D.
Alyssa Guthrie, M.S.Sc.
Melanie Muise, M.A.
Ben Vandermeer, M.Sc.
Walie M. Aktary, B.Sc., B.Ed.
Dion Pasichnyk, B.Sc.
Jennifer C. Seida, M.P.H.

Clinical Investigator:
Lois Donovan, M.D., FRCPC

AHRQ Publication No. 12(13)-E021-EF
October 2012

This report is based on research conducted by the University of Alberta Evidence-based Practice Center under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2007-10021-I). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. No statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact EffectiveHealthCare@ahrq.hhs.gov.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

Suggested citation: Hartling L, Dryden DM, Guthrie A, Muise M, Vandermeer B, Aktary WM, Pasichnyk D, Seida JC, Donovan L. Screening and Diagnosing Gestational Diabetes Mellitus. Evidence Report/Technology Assessment No. 210. (Prepared by the University of Alberta Evidence-based Practice Center under Contract No. 290-2007-10021-I.) AHRQ Publication No. 12(13)-E021-EF. Rockville, MD: Agency for Healthcare Research and Quality. October 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions, and new health care technologies and strategies.

The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments. To bring the broadest range of experts into the development of evidence reports and health technology assessments, AHRQ encourages the EPCs to form partnerships and enter into collaborations with other medical and research organizations. The EPCs work with these partner organizations to ensure that the evidence reports and technology assessments they produce will become building blocks for health care quality improvement projects throughout the Nation. The reports undergo peer review and public comment prior to their release as a final report.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality.

We welcome comments on this evidence report. Comments may be sent by mail to the Task Order Officer named in this report to: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.hhs.gov.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director
Evidence-based Practice Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Suchitra Iyer, Ph.D.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project: Tamara Durec (searching), Andrea Milne (searching, technical support), Teodora Radisic (article retrieval), Jocelyn Shulhan (screening), Annabritt Chisholm (research support, reference management), and Noah Toppings (development of Figure 1). We thank Dr. Alun Edwards for providing clinical input throughout the project.

Technical Expert Panel

Howard Berger, M.D.
Assistant Professor
Health Policy, Management, and Evaluation
University of Toronto
Toronto, ON

Siri L. Kjos, M.D., M.S.Ed.
Clinical Faculty
David Geffen School of Medicine
University of California, Los Angeles
Torrance, CA

Joy Melnikow, M.D., M.P.H
Professor
Family and Community Medicine
University of California, Davis
Sacramento, CA

Wanda K. Nicholson, M.D., M.P.H., M.B.A.
Associate Professor
Obstetrics and Gynecology
University of North Carolina
Chapel Hill, NC

Leann Olansky, M.D.
Medical Director
Lennon Diabetes Center, Huron Hospital
Cleveland Clinic East
Cleveland, OH

Peer Reviewers

Florence Brown, M.D.
Assistant Professor, Harvard Medical School
Director, Joslin-Beth Israel Deaconess Diabetes and Pregnancy Program
Boston, MA

William Callaghan, M.D., M.P.H.
Senior Scientist
Maternal and Infant Health Branch
Division of Reproductive Health
National Center for Chronic Disease Prevention and Health Promotion
Atlanta, GA

Robert P. Kauffman, M.D.
Professor
Department of Obstetrics and Gynecology
Texas Tech University
Lubbock, TX

Hélène Long, M.D.
Endocrinologist
Cité de la Santé de Laval Hospital
Laval, QC, Canada

Julia Lowe, M.D.
Associate Professor, Endocrinology
University of Toronto, Sunnybrook Health Sciences Center
Toronto, ON, Canada

James Peter VanDorsten, M.D.
Professor
Department of Obstetrics and Gynecology
Medical University of South Carolina
Charleston, SC

Screening and Diagnosing Gestational Diabetes Mellitus

Structured Abstract

Background. There is uncertainty as to the optimal approach for screening and diagnosis of gestational diabetes mellitus (GDM). Based on systematic reviews published in 2003 and 2008, the U.S. Preventive Services Task Force concluded that there was insufficient evidence upon which to make a recommendation regarding routine screening of all pregnant women.

Objectives. (1) Identify properties of screening tests for GDM, (2) evaluate benefits and harms of screening for GDM, (3) assess the effects of different screening and diagnostic thresholds on outcomes for mothers and their offspring, and (4) determine the benefits and harms of treatment for a diagnosis of GDM.

Data Sources. We searched 15 electronic databases from 1995 to May 2012, including MEDLINE and Cochrane Central Register of Controlled Trials (which contains the Cochrane Pregnancy and Childbirth Group registry); gray literature; Web sites of relevant organizations; trial registries; and reference lists.

Methods. Two reviewers independently conducted study selection and quality assessment. One reviewer extracted data, and a second reviewer verified the data. We included published randomized and nonrandomized controlled trials and prospective and retrospective cohort studies that compared any screening or diagnostic test with any other screening or diagnostic test; any screening with no screening; women who met various thresholds for GDM with those who did not meet various criteria, where women in both groups did not receive treatment; any treatment for GDM with no treatment. We conducted a descriptive analysis for all studies and meta-analyses when appropriate. Key outcomes included preeclampsia, maternal weight gain, birth injury, shoulder dystocia, neonatal hypoglycemia, macrosomia, and long-term metabolic outcomes for the child and mother.

Results. The search identified 14,398 citations and included 97 studies (6 randomized controlled trials, 63 prospective cohort studies, and 28 retrospective cohort studies).

Prevalence of GDM varied across studies and diagnostic criteria: American Diabetes Association (75 g) 2 to 19 percent; Carpenter and Coustan 3.6 to 38 percent; National Diabetes Data Group 1.4 to 50 percent; and World Health Organization 2 to 24.5 percent. Lack of a gold standard for the diagnosis of GDM and little evidence about the accuracy of screening strategies for GDM remain problematic. The 50 g oral glucose challenge test with a glucose threshold of 130 mg/dL versus 140 mg/dL improves sensitivity and reduces specificity. Both thresholds have high negative predictive values (NPV) but variable positive predictive values (PPVs) across a range of prevalence. There was limited evidence for the screening of GDM diagnosed less than 24 weeks' gestation (three studies). One study compared the International Association of Diabetes in Pregnancy Study Groups' (IADPSG) diagnostic criteria with a two-step strategy. Sensitivity was 82 percent, specificity was 94 percent.

Only two studies examined the effects on health outcomes from screening for GDM. One retrospective cohort study (n=1,000) showed more cesarean deliveries in the screened group. A survey within a prospective cohort study (n=93) found the same incidence of macrosomia (≥4.3 kg) in screened and unscreened groups (7 percent each group).

Thirty-eight studies examined health outcomes for women who met different criteria for GDM and did not undergo treatment. Methodologically strong studies showed a continuous positive relationship between increasing glucose levels and the incidence of primary cesarean section and macrosomia. One of these studies also found significantly fewer cases of preeclampsia, cesarean section, shoulder dystocia and/or birth injury, clinical neonatal hypoglycemia, and hyperbilirubinemia for women without GDM compared with those meeting IADPSG criteria. Among the other studies, fewer cases of preeclampsia were observed for women with no GDM and women who were false positive versus those meeting Carpenter and Coustan criteria. For maternal weight gain, few comparisons showed differences. For fetal birth trauma, single studies showed no differences for women with Carpenter and Coustan GDM and World Health Organization impaired glucose tolerance versus women without GDM. Women diagnosed based on National Diabetes Data Group GDM had more fetal birth trauma compared with women without GDM. Fewer cases of macrosomia were seen in the group without GDM compared with Carpenter and Coustan GDM, Carpenter and Coustan 1 abnormal oral glucose tolerance test, National Diabetes Data Group GDM, National Diabetes Data Group false positives, and World Health Organization impaired glucose tolerance. Fewer cases of neonatal hypoglycemia were found among patient groups without GDM compared with those meeting Carpenter and Coustan criteria. There was more childhood obesity for Carpenter and Coustan GDM versus patient groups with no GDM.

Eleven studies compared diet modification, glucose monitoring, and insulin as needed with no treatment. Moderate evidence showed fewer cases of preeclampsia in the treated group. The evidence was insufficient for maternal weight gain and birth injury. Moderate evidence found less shoulder dystocia with treatment for GDM. Low evidence showed no difference for neonatal hypoglycemia between treated and untreated GDM. Moderate evidence showed benefits of treatment for reduction of macrosomia (>4,000 g). There was insufficient evidence for long-term metabolic outcomes among offspring.

Five studies provided data on harms of treating GDM. No difference was found for cesarean delivery, induction of labor, small for gestational age, or admission to a neonatal intensive care unit. There were significantly more prenatal visits among those treated.

Conclusions. While evidence supports a positive association with increasing plasma glucose on a 75 g or 100 g oral glucose tolerance test and macrosomia and primary cesarean section, clear thresholds for increased risk were not found. The 50 g oral glucose challenge test has high NPV but variable PPV. Treatment of GDM results in less preeclampsia and macrosomia. Current evidence does not show that treatment of GDM has an effect on neonatal hypoglycemia or future poor metabolic outcomes. There is little evidence of short-term harm from treating GDM other than an increased demand for services. Research is needed on the long-term metabolic outcome for offspring as a result of GDM and its treatment, and the “real world” effects of GDM treatment on use of care.

Contents

Tables

Figures

Appendixes

Executive Summary

Introduction

Gestational Diabetes Mellitus

Gestational diabetes mellitus (GDM) is defined as glucose intolerance first discovered in pregnancy. Pregestational diabetes mellitus refers to any type of diabetes diagnosed before pregnancy. Pregnant women with pregestational diabetes experience an increased risk of poor maternal, fetal, and neonatal outcomes.[1] The extent to which GDM predicts adverse outcomes for mother, fetus, and neonate is less clear.

Depending on the diagnostic criteria used and the population screened, the prevalence of GDM ranges from 1.1 to 25.5 percent of pregnancies in the United States.[2-4] In 2009, the Centers for Disease Control and Prevention reported a prevalence of 4.8 percent of diabetes in pregnancy. An estimated 0.5 percent of these cases likely represented women with pregestational diabetes. Data from the international Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study[3] indicate that 6.7 percent of the women met a fasting plasma glucose threshold of 95 mg/dL (5.3 mmol/L), which is in keeping with the Carpenter and Coustan[5] (CC) criteria that are in common practice in North America. In contrast, 17.8 percent of women were diagnosed with GDM using the International Association of the Diabetes in Pregnancy Study Groups (IADPSG) criteria in which lower glucose thresholds diagnose GDM.

The prevalence of GDM is not only influenced by diagnostic criteria but also by population characteristics. In a recent publication, data from the Hyperglycemia and Adverse Pregnancy Outcome Study (HAPO) demonstrated wide variability in GDM prevalence across a number of study centers, both internationally and within the United States, even when the same diagnostic criteria are applied (i.e., the IADPSG criteria).[6] Prevalence in the United States ranged from 15.5 percent in Providence, RI, to 25.5 percent in Bellflower, CA. There are ethnic differences in the prevalence of GDM in the United States. Native Americans, Asians, Hispanics, and African-American women are at higher risk than non-Hispanic white women.[7] Data from 2000 showed that prevalence was highest among Asian and Hispanic women (~7 to 8 percent), intermediate among African-American women (~6 percent), and lower among non-Hispanic white women (~5 percent) based on CC criteria and/or hospital discharge diagnosis.[7] The rate of increase of prevalence over the past 10 years has been highest for Asian and African-American women.[7]

The incidence of GDM has increased over the past decades in parallel with the increase in rates of obesity and type 2 diabetes mellitus, and this trend is expected to continue.[8] It is unclear how much the increase in obesity will affect the proportion of women diagnosed with overt diabetes during pregnancy versus transient pregnancy-induced glucose intolerance.

GDM is usually diagnosed after 20 weeks' gestation when placental hormones that have the opposite effect of insulin on glucose metabolism increase substantially. Women with adequate insulin secreting capacity overcome this insulin resistance of pregnancy by secreting more endogenous insulin to maintain normal blood glucose. Women with less adequate pancreatic reserve are unable to produce sufficient insulin to overcome the increase in insulin resistance, and glucose intolerance results.

Glucose abnormalities in women with GDM usually resolve postpartum, but commonly recur in subsequent pregnancies. Women with GDM have an increased risk of future development of overt diabetes. The cumulative incidence of diabetes after a diagnosis of GDM varies widely

depending on maternal body mass index (BMI), ethnicity, and time since index pregnancy, and it may reach levels as high as 60 percent.[9] When glucose abnormalities persist postpartum in a woman with GDM, her diabetes is recategorized as overt diabetes. When this occurs, the likelihood that this woman had pregestational (i.e., overt) diabetes increases, especially if the diagnosis of GDM occurred before 20 weeks' gestation and glucose levels were markedly elevated in pregnancy.

Studies investigating pregnancy outcomes of women with GDM show considerable variability in the proportion of women with suspected pregestational diabetes. This variability contributes to the confusion surrounding the true morbidity of GDM. In an attempt to enable better comparability across future studies and more accurate risk stratification of pregnant women with diabetes, recommendations[10] have proposed that women with more severe glucose abnormalities in pregnancy be excluded from the diagnosis of GDM. The expectation is that this would exclude women with overt diabetes from the population of women defined as having GDM. This proposal is in contrast to the older definition of GDM, which includes any degree of glucose intolerance first discovered in pregnancy.

Risk Factors

Risk factors for GDM include greater maternal age, higher BMI, member of an ethnic group at increased risk for development of type 2 diabetes mellitus (i.e., Hispanic, African, Native American, South or East Asian, or Pacific Islands ancestry), polyhydramnios, past history of GDM, macrosomia in a previous pregnancy, history of unexplained stillbirth, type 2 diabetes mellitus in a first degree relative, polycystic ovary syndrome, and metabolic syndrome.[11] Low risk of GDM is usually defined as young (age less than 25 or 30 years), non-Hispanic white, normal BMI (25 kg/m^2 or less), no history of previous glucose intolerance or adverse pregnancy outcomes associated with GDM, and no first degree relative with known diabetes.[7,12] Women at high risk of GDM are usually defined as having two or more risk factors for GDM. Women at moderate risk of GDM do not satisfy all criteria of women at low risk, but they lack two or more risk factors for GDM.

Screening and Diagnostic Strategies

The 2008 U.S. Preventive Services Task Force (USPSTF) evidence review on screening for GDM concluded that at that time, "evidence was insufficient to assess the balance of benefits and harms of screening for GDM either before or after 24 weeks' gestation."[13] The report suggested that "…until there was better evidence, clinicians should discuss screening for GDM with their patient and make case-by-case decisions. Discussions should include information about the uncertainty of benefits and harms as well as the frequency of positive screening test results."

The 2001 practice guidelines of the American College of Obstetricians and Gynecologists (ACOG) endorsed risk factor-based screening for GDM, recognizing that low-risk women may be less likely to benefit from screening with glucose measurements. Women were considered low risk of GDM if they met all the following criteria: (1) younger than 25 years; (2) not a member of an ethnic group at high risk for development of type 2 diabetes mellitus; (3) BMI of 25 kg/m^2 or less; (4) no history of previous glucose intolerance or adverse pregnancy outcomes associated with GDM; and (5) no first degree relative with known diabetes. ACOG plans to update its 2001 practice guidelines on GDM based on the proceedings of the 2012 National Institutes of Health consensus conference on GDM diagnosis. Until 2011, the American Diabetes Association (ADA) also endorsed no screening for pregnant woman who met all the criteria

mentioned above for low risk of GDM. In 2011 the ADA changed their recommendations to endorse glucose testing for GDM in all pregnant women who do not have a diagnosis of pregestational diabetes.

Common practices of glucose screening for GDM in North America involve a two-step approach in which patients with abnormal results on a screening test receive a subsequent diagnostic test.[14] Typically, a 50 g oral glucose challenge test (OGCT) is initially administered between 24 and 28 weeks' gestation in a nonfasting state, in women at moderate risk (i.e., women who do not meet all low risk criteria but lack two or more risk factors for GDM). The test is administered earlier in gestation for women at high risk of GDM (i.e., multiple risk factors for GDM) and repeated at 24–28 weeks' gestation if initial surveillance is normal. Patients who meet or exceed a screening threshold (usually 130 mg/dL or 140 mg/dL) receive a more involved diagnostic test—the oral glucose tolerance test (OGTT), in which a 75 g or 100 g oral glucose load is administered in a fasting state, and plasma glucose levels are evaluated after 1, 2, or 3 hours. A diagnosis of GDM is made in pregnant women when one or more glucose values fall at or above the specified glucose thresholds. Alternatively, a one-step method in which all patients or high-risk patients forego the screening test and proceed directly to the OGTT has been recommended.[15]

The absence of a universally accepted gold standard for the diagnosis of GDM has resulted in a variety of recommended diagnostic glucose thresholds that have been endorsed by different stakeholders (Table A). These criteria reflect changes that have occurred in laboratory glucose measurements over the years and in new evidence that suggests the ability of different glucose thresholds to predict poor pregnancy outcomes. The different diagnostic criteria and thresholds result in different estimates of the prevalence of GDM.

In 2004, a cross-sectional study reported that universal screening was the most common practice in the United States, with 96 percent of obstetricians routinely screening for GDM.[16] In contrast, the guidelines of ACOG and the ADA at that time stated that women at low risk for GDM were unlikely to benefit from screening.[14,17] Since only 10 percent of pregnant women were categorized as low risk, some argued that selective screening contributed to confusion, with little benefit and potential for harm.[18] Of particular concern was the association between risk factor-based screening and high rates of false negative results.[19] Others have endorsed alternative risk scoring systems for screening.[20]

The IADPSG, an international consensus group with representation from multiple obstetrical and diabetes organizations, recently spearheaded a reexamination of the definition of GDM in an attempt to bring uniformity to GDM diagnoses.[21] The IADPSG recommended that a one-step 75 g OGTT be given to all pregnant women who do not have a diagnosis of overt diabetes. They also recommended that a single glucose value, rather than at least two abnormal values at or above diagnostic glucose thresholds on the OGTT be accepted as sufficient for a diagnosis of GDM. The diagnostic glucose thresholds recommended by the IADPSG were the maternal glucose values from the HAPO study[3] that identified a 1.75-fold increase (adjusted odds ratio relative to the mean cohort glucose values) in large for gestational age, elevated C-peptide, high neonatal body fat, or in a combination of these factors. Since overt diabetes is often asymptomatic, may not have been screened for before conception, has a prevalence that is increasing dramatically in reproductive-age women, and carries a higher risk for poor pregnancy outcomes,[22] the IADPSG also recommended that all women, or at least women from high-risk groups for type 2 diabetes mellitus, be screened for overt diabetes at their first prenatal visit and excluded from the diagnosis of GDM using one of the following criteria: fasting plasma glucose

≥126 mg/dL (7.0 mmol/L), glycated hemoglobin (HbA1c) ≥6.5 percent (Diabetes Chronic Complications Trial/United Kingdom Prospective Diabetes Study standardized), or a random plasma glucose ≥200 mg/dL (11.1 mmol/L) confirmed by one of the first two measures.

Treatment Strategies

Initial treatment for GDM involves diet modification, glucose monitoring, and moderate exercise. When dietary management does not achieve desired glucose control, insulin or oral antidiabetic medications may be used.[23] Increased prenatal surveillance may also occur as well as changes in delivery management depending on fetal size and the effectiveness of measures to control glucose.

Scope of the Review

Based on systematic reviews published in 2003 and 2008, the USPSTF concluded that there was insufficient evidence upon which to make a recommendation regarding routine screening of all pregnant women for GDM.[13,24] Several key studies have been published since the 2008 USPSTF evidence report.[3,8,25] The National Institutes of Health's Office of Medical Applications of Research (OMAR) commissioned this report (specifically Key Questions 3 to 5, see section below), which the Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Center (EPC) Program conducted. OMAR will use the review to inform members of consensus meetings and inform guideline development. The USPSTF joined this effort and will use the review to update its recommendation on screening for GDM (Key Questions 1 and 2).

The primary aims of this review were to (1) identify the test properties of screening and diagnostic tests for GDM, (2) evaluate the potential benefits and harms of screening at ≥24 weeks and <24 weeks' gestation, (3) assess the effects of different screening and diagnostic thresholds on outcomes for mothers and their offspring, and (4) determine the effects of treatment in modifying outcomes for women diagnosed with GDM. The benefits and harms of treatments were considered in this review to determine the downstream effects of screening on health outcomes. The intent of this review was also to assess whether evidence gaps in the previous USPSTF reviews have been filled. These gaps included lack of sufficient evidence to determine whether maternal or fetal complications are reduced by screening; lack of screening studies with adequate power to evaluate health outcomes such as mortality, neonatal intensive care unit (NICU) admissions, hyperbilirubinemia; limited evidence on the accuracy of screening strategies; and insufficient evidence on the benefits of treating GDM in improving health outcomes.

Table A. Diagnostic criteria and plasma glucose thresholds for gestational diabetes mellitus

Organization	Year	Testing Schedule	Abnormal Value(s)	Threshold (Equal to or Greater Than)			
				0 (h)	1 (h)	2 (h)	3 (h)
ADA	1999[26]	50 g OGCT	1	—	140 mg/dL 7.8 mmol/L	—	—
		100 g OGTT	2 or more	105 mg/dL 5.8 mmol/L	190 mg/dL 10.5 mmol/L	165 mg/dL 9.1 mmol/L	145 mg/dL 8.0 mmol/L
ADA Low risk† excluded	2000-2010[10,27-36]	50 g OGCT	1	—	130 mg/dL 7.2 mmol/L or 140 mg/dL 7.8 mmol/L	—	—
		100 g or 75 g OGTT after overnight fast ≥8hr	2 or more	95 mg/dL 5.3 mmol/L	180 mg/dL 10.0 mmol/L	155 mg/dL 8.6 mmol/L	140 mg/dL 7.8 mmol/L (3 hr value only for 100 g test)
IADPSG ADA	2011[37]	75 g OGTT	1 or more	92 mg/dL 5.1 mmol/L	180 mg/dL 10.0 mmol/L	153 mg/dL 8.5 mmol/L	—
1. CC 2. 4th IWC (same) 3. 5th IWC (same as 4th but 75 g accepted with same glucose thresholds)	1. 1982[5] 2. 1998[38] 3. 2007[39]	50 g OGCT	1	—	130 mg/dL 7.2 mmol/L	—	—
		100 g OGTT	2 or more	95 mg/dL 5.3 mmol/L	180 mg/dL 10.0 mmol/L	155 mg/dL 8.6 mmol/L	140 mg/dL 7.8 mmol/L
NDDG	1979[40]	50 g OGCT	—	—	—	—	—
		100 g OGTT	2 or more	105 mg/dL 5.8 mmol/L	190 mg/dL 10.5 mmol/L	165 mg/dL 9.1 mmol/L	145 mg/dL 8.0 mmol/L
WHO	1999 WHO consultation[41]	75 g OGTT	1	6.1 mmol/L for IGT of pregnancy; 7.0 mmol/L for Dx of DM	—	140 mg/dL 7.8 mmol/L for IGT of pregnancy; 200 mg/dL 11.1 mmol/L for Dx of DM	—
WHO	1985 WHO study group report[42]	75 g OGTT	1	7.8 mmol/L 140 mg/dL for IGT of pregnancy	—	7.8 mmol/L (140 mg/dL); for IGT of pregnancy; 200 (11.1 mmol/L) for Dx of DM	—

Table A. Diagnostic criteria and plasma glucose thresholds for gestational diabetes mellitus (continued)

Organization	Year	Testing Schedule	Abnormal Value(s)	Threshold (Equal to or Greater Than)			
				0 (h)	1 (h)	2 (h)	3 (h)
CDA	2003, 2008[43,44]	50 g OGCT	1	—	140 mg/dL 7.8 mmol/L or 186 mg/dL, 10.3 mmol/L Dx GDM	—	—
		75 g	2 or more	95 mg/dL 5.3 mmol/L	191 mg/dL 10.6 mmol/L	160 mg/dL 8.9 mmol/L	—
ACOG – risk factor 4th IWC	2001[14,45]	50 g	1	—	130 mg/dL 7.2 mmol/L or 140 mg/dL 7.8 mmol/L	—	—
		100 g CC	2 or more	95 mg/dL 5.3 mmol/L	180 mg/dL 10.0 mmol/L	155 mg/dL 8.5 mmol/L	140 mg/dL 7.8 mmol/L
		100 g NDDG	2 or more	105 mg/dL 5.8 mmol/L	190 mg/dL 10.5 mmol/L	165 mg/dL 9.1 mmol/L	145 mg/dL 8.0 mmol/L
3rd IWC	1991[46]	100 g OGTT	2 or more	105 mg/dL 5.8 mmol/L	190 mg/dL 10.5 mmol/L	165 mg/dL 9.1 mmol/L	145 mg/dL 8.0 mmol/L
ADIPS	1998[47]	50 g or 75 g nonfasting	1	—	140 mg/dL 7.8 mmol/L (50 g) or 144 mg/dL 8.0 mmol/L (75 g)	—	—
		75 g fasting	1	99 mg/dL 5.5 mmol/L	—	144 mg/dL 8.0 mmol/L or 162 mg/dL 9.0 mmol/L*	—

Table A. Diagnostic criteria and plasma glucose thresholds for gestational diabetes mellitus (continued)

Organization	Year	Testing Schedule	Abnormal Value(s)	Threshold (Equal to or Greater Than)			
				0 (h)	1 (h)	2 (h)	3 (h)
EASD	1996[48]	75 g	1	108 mg/dL 6.0 mmol/L	—	162 mg/dL 9.0 mmol/L	—
USPSTF (Grade 1 recommendation)	2008‡	Risk assessment 50 g OGCT	1	—	130 mg/dL 7.2 mmol/L or 140 mg/dL 7.8 mmol/L	—	—
		100 g OGTT	2 or more	NR	NR	NR	NR

ACOG = American College of Obstetricians and Gynecologists; ADA = American Diabetes Association; ADIPS = Australasian Diabetes in Pregnancy Society; CC = Carpenter, Coustan; CDA = Canadian Diabetes Association; DM = diabetes mellitus; Dx = diagnosis; EASD = European Association for the Study of Diabetes; GDM = gestational diabetes mellitus; IADPSG = International Association of Diabetes in Pregnancy Study Groups; IGT = impaired glucose tolerance; IWC = International Workshop Conference; NDDG = National Diabetes Data Group; NR = not reported; OGCT = oral glucose challenge test; OGTT = oral glucose tolerance test; USPSTF = U.S. Preventive Services Task Force; WHO = World Health Organization

†Low risk defined as age <25 yr, normal body weight, no first degree relative with DM, no history of abnormal glucose, no history of poor obstetrical outcomes, not of high risk ethnicity for DM.

*in New Zealand.

‡ Screening for GDM: USPSTF recommendation statement Ann Intern Med 2008;148(10):759-65.

Key Questions

OMAR and USPSTF developed the Key Questions for this evidence synthesis to inform members of consensus meetings and inform guideline development; OMAR specifically developed Key Questions 3 to 5. Investigators from the University of Alberta EPC worked in consultation with representatives from the AHRQ EPC Program, OMAR and the USPSTF, and a panel of Technical Experts to operationalize the Key Questions. The Technical Expert Panel provided content and methodological expertise throughout the development of this evidence synthesis. Participants in this panel are identified in the front matter of this report. The Key Questions are as follows:

Key Question 1: What are the sensitivities, specificities, reliabilities, and yields of current screening tests for GDM? (a) After 24 weeks' gestation? (b) During the first trimester and up to 24 weeks' gestation?

Key Question 2: What is the direct evidence on the benefits and harms of screening women (before and after 24 weeks' gestation) for GDM to reduce maternal, fetal, and infant morbidity and mortality?

Key Question 3: In the absence of treatment, how do health outcomes of mothers who meet various criteria for GDM and their offspring compare to those who do not meet the various criteria?

Key Question 4: Does treatment modify the health outcomes of mothers who meet various criteria for GDM and their offspring?

Key Question 5: What are the harms of treating GDM and do they vary by diagnostic approach?

Methods

Literature Search

We systematically searched the following bibliographic databases for studies published from 1995 to May 2012: MEDLINE® Ovid, Ovid MEDLINE® In-Process & Other Non-Indexed Citations, Cochrane Central Register of Controlled Trials (contains the Cochrane Pregnancy and Childbirth Group, which hand searches journals pertinent to its content area and adds relevant trials to the registry), Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts of Reviews of Effects (DARE), Global Health, Embase, Pascal CINAHL Plus with Full Text (EBSCO host), BIOSIS Previews® (Web of KnowledgeSM), Science Citation Index Expanded® and Conference Proceedings Citation Index- Science (both via Web of ScienceSM), PubMed®, LILACS (Latin American and Caribbean Health Science Literature), National Library of Medicine (NLM) Gateway, and OCLC ProceedingsFirst and PapersFirst. We searched trial registries, including the WHO International Clinical Trials Registry Platform (ICTRP), ClinicalTrials.gov, and Current Controlled Trials. We limited the search to trials and cohort studies published in English.

We searched the Web sites of relevant professional associations and research groups, including the ADA, IADPSG, International Symposium of Diabetes in Pregnancy, and Diabetes

in Pregnancy Society for conference abstracts and proceedings from the past 3 years. We reviewed the reference lists of relevant reviews (including the 2008 USPSTF review) and studies that were included in this report.

Study Selection

Two reviewers independently screened the titles and abstracts using broad inclusion criteria. We retrieved the full text of articles classified as "include" or "unclear." Two reviewers independently assessed each full-text article using a priori inclusion criteria and a standardized form. We resolved disagreements by consensus or third-party adjudication.

We included published randomized controlled trials (RCTs), nonrandomized controlled trials (NRCTs), and prospective and retrospective cohort studies. For Key Question 1, we excluded retrospective cohort studies. We included studies of pregnant women ≥24 weeks' gestation or <24 weeks' gestation, with no known history of preexisting diabetes. Comparisons of interest varied by Key Question and were as follows: Key Question 1 – any GDM screening or diagnostic test compared with any GDM reference standard or other screening or diagnostic test; Key Question 2 – any GDM screening versus no GDM screening; Key Question 3 – women who met various thresholds for GDM versus those who did not meet various criteria for GDM, where women in both groups did not receive treatment; Key Questions 4 and 5 – any treatment for GDM, including but not limited to dietary advice, blood glucose monitoring, insulin therapy (all preparations), and oral hypoglycemic agents versus no treatment. Studies meeting these eligibility criteria were included if they reported data for at least one outcome specified in the Key Questions. We included studies regardless of setting and duration of followup.

Quality Assessment

Two reviewers independently assessed the methodological quality of studies and resolved discrepancies by discussion and consensus. For Key Question 1, we used the QUADAS-2 checklist[49] to assessthe quality of diagnostic accuracy studies. We assessed the internal validity of RCTs and NRCTs using the Cochrane Collaboration Risk of Bias tool. For cohort studies, we used the Newcastle-Ottawa Scale. For Key Questions 2 to 5, we summarized the quality of individual studies as "good," "fair," or "poor" based on criteria specific to each tool.

Data Extraction and Synthesis

One reviewer extracted data using a standardized form, and a second reviewer checked the data for accuracy and completeness. We extracted information on study characteristics, inclusion and exclusion criteria, participant characteristics, details of the interventions or diagnostic/screening tests (as appropriate), and outcomes. Reviewers resolved discrepancies by consensus or in consultation with a third party.

For each Key Question, we presented evidence tables detailing each study and provided a qualitative description of results. For Key Question 1, we constructed 2x2 tables and calculated sensitivity, specificity, positive and negative predictive values, reliability (i.e., accuracy), and yield (i.e., prevalence) of the screening or diagnostic tests. If studies were clinically homogenous, we pooled sensitivities and specificities using a hierarchical summary receiver-operator curve and bivariate analysis of sensitivity and specificity.[50] For the other Key Questions, we combined studies in a meta-analysis if the study design, population, comparisons, and outcomes were sufficiently similar. Results were combined using random effects models.

We quantified statistical heterogeneity using the I-squared (I^2) statistic. When I^2 was greater than 75 percent, we did not pool results, and we investigated potential sources of heterogeneity.

Strength of the Body of Evidence

Two independent reviewers graded the strength of the evidence for Key Questions 3 and 4 using the EPC GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach and resolved discrepancies by discussion and consensus. We graded the evidence for the following key outcomes: birth injury, preeclampsia, neonatal hypoglycemia, maternal weight gain, and long-term metabolic outcomes of the child and mother. We made a post hoc decision to grade shoulder dystocia and macrosomia. These were not included in the protocol as outcomes that would be graded but were felt by the clinical investigators to be important to grade during the course of preparing the review. For each outcome, we assessed four major domains: risk of bias (rated as low, moderate, or high), consistency (rated as consistent, inconsistent, or unknown), directness (rated as direct or indirect), and precision (rated as precise or imprecise). The overall strength of evidence was graded as high, moderate, low, or insufficient.

Applicability

We assessed the applicability of the body of evidence following the PICOTS (population, intervention, comparator, outcomes, timing of outcome measurement, and setting) format used to assess study characteristics. Factors that may potentially limit applicability were discussed.

Peer Review and Public Commentary

Peer reviewers were invited to provide written comments on the draft report based on their clinical, content, or methodologic expertise. Peer reviewer comments on the draft report were addressed by the EPC in preparation of the final draft of the report. Peer reviewers do not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments are documented and will be published 3 months after the publication of the Evidence Report.

Potential reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Invited peer reviewers may not have any financial conflict of interest greater than $10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through AHRQ's public comment mechanism.

Results

Description of Included Studies

The search identified 14,398 citations, and 97 studies were included: 6 RCTs, 63 prospective cohort studies, and 28 retrospective cohort studies. The studies were published between 1995 and 2012 (median 2004). Studies were conducted in the United States (24 percent), Europe (23 percent), Asia (22 percent), the Middle East (20 percent), Australia (4 percent), Central and South America (3 percent), and Canada (4 percent). The number of women enrolled in each study ranged from 32 to 23,316 (median 750). The mean age of study participants was 30 years.

Forty-eight studies (50 percent) analyzed women tested for GDM between 24 and 28 weeks, with an OGCT taking place first and the OGTT following within 7 days. Thirty-one studies (32 percent) did not specify when screening or diagnostic procedures took place. Eighteen studies (18 percent) screened or tested within unique time ranges. Of these, one study screened participants with an OGCT at 21–23 weeks followed by a diagnostic OGTT at 24–28 weeks; another screened a group of participants after 37 weeks; one study screened before 24 weeks; another screened women at risk between 14 and 16 weeks, with normal women screened at the usual 24–28 weeks; and one study screened between 16 and 20 weeks or between 17 and 21 weeks followed by a diagnostic OGTT at 26–32 weeks. Remaining studies generally provided broader screening times ranging from 21 to 32 weeks' gestation. Studies employing WHO criteria generally screened further into gestation as only an OGTT was performed: one study screened at 28–32 weeks, and another study screened women at high risk at 18–20 weeks and others at 28–30 weeks.

Methodological Quality of Included Studies

The methodological quality was assessed using different tools depending on the Key Question and study design: QUADAS-2 was used for Key Question 1; for Key Questions 2 to 5, the Cochrane Risk of Bias tool was used for RCTs and the Newcastle Ottawa Scale was used for cohort studies. The methodological quality of studies is summarized for each Key Question below.

Results of Included Studies

The results are presented by Key Question in the sections that follow. A summary of the results for all Key Questions is provided in Table D at the end of the Executive Summary.

Key Question 1

Fifty-one studies provided data for Key Question 1, which examined the diagnostic test characteristics and prevalence of current screening and diagnostic tests for GDM. Studies were conducted in a range of geographic regions: 11 in North America, 10 in Europe, 12 in Asia, 15 in the Middle East, 2 in South America, and 1 in Australia. Studies reported on findings for a number of screening tests, including the 50 g OGCT, fasting plasma glucose (FPG), and risk factor-based screening, as well as other, less common tests such as HbA1c, serum fructosamine, and adiponectin. GDM was confirmed using criteria developed by different groups, including CC, ADA, National Diabetes Data Group (NDDG), and WHO. The lack of a gold standard to confirm a diagnosis of GDM limits the ability to compare the results of studies that have used different diagnostic criteria. Different criteria result in different rates of prevalence, regardless of similarities across study settings and patient characteristics. A summary of the results is provided in Table D.

Methodological quality of the studies was assessed using the QUADAS-2 tool. The domain of patient selection was rated as low risk for 53 percent and unclear risk for 22 percent of the studies. Overall, 55 percent were assessed as having high concerns about applicability for this domain. This was primarily because these studies were conducted in developing countries and used the WHO criteria to diagnose GDM. The domain of the index test was generally rated as low risk of bias (53 percent). Concern about applicability was assessed as low (82 percent). The domain of the reference standard (i.e., the criteria used to confirm a diagnosis of GDM) was rated as high or unclear risk (80 percent). For most studies, the result of the screening test was

used to determine whether patients underwent further testing for GDM (lack of blinding) or it was unclear. Concern about applicability for this domain was assessed as low (84 percent). The domain of flow and timing was assessed as low risk of bias in 39 percent of studies. However, 35 percent were assessed as unclear risk of bias because not all patients received a confirmatory reference standard if the screening test was below a certain threshold, so there is a risk of diagnostic review bias.

Nine studies provided data to estimate sensitivity and specificity of a 50 g OGCT (cutoff ≥140 mg/dL); GDM was confirmed using a 100 g, 3-hour OGTT using CC criteria. Sensitivity and specificity were 85 percent (95% CI, 76 to 90) and 86 percent (95% CI, 80 to 90), respectively. Prevalence ranged from 3.8 to 31.9 percent. When prevalence was less than 10 percent, PPV ranged from 18 to 27 percent; when prevalence was 10 percent or more, PPV ranged from 32 to 83 percent. The median NPV for all studies was 98 percent.

Six studies reported results for a 50 g OGCT (cutoff ≥130 mg/dL); GDM was confirmed using the CC criteria. Sensitivity was 99 percent (95% CI, 95 to 100) and specificity was 77 percent (95% CI, 68 to 83). Prevalence ranged from 4.3 to 29.8 percent. When prevalence was less than 10 percent, PPV ranged from 11 to 27 percent; when prevalence was 10 percent or more, PPV ranged from 31 to 62 percent. The median NPV for all studies was 100 percent.

One study assessed a 50 g OGCT with a cutoff value of ≥200 mg/dL; GDM was confirmed using the CC criteria. Prevalence was 6.4 percent. Sensitivity, specificity, PPV and NPV were all 100 percent.

The evidence showed that the 50 g OGCT with the 130 mg/dL cutpoint had higher sensitivity when compared with the 140 mg/dL cutpoint; however, specificity was lower. Both thresholds have high NPVs, but variable PPVs across a range of GDM prevalence. The Toronto Trihospital study found evidence to support the use of the lower screening cutpoint for higher risk patients, and the higher screening cutpoint for lower risk patients.[12]

Seven studies assessed a 50 g OGCT (≥140 mg/dL); GDM was confirmed using the NDDG criteria. Sensitivity was 85 percent (95% CI, 73 to 92) and specificity was 83 percent (95% CI, 78 to 87). Prevalence ranged from 1.4 to 45.8 percent. When prevalence was less than 10 percent, PPV ranged from 12 to 39 percent; prevalence was more than 10 percent in one study and PPV was 57 percent. The median NPV for all studies was 99 percent. Three studies that assessed a 50 g OGCT (≥130 mg/dL) using NDDG were not pooled. Prevalence ranged from 16.7 to 35.3 percent. PPV ranged from 20 to 75 percent; NPV ranged from 86 to 95 percent.

Three studies assessed a 50 g OGCT (different thresholds); GDM was confirmed using the ADA 2000-2010 75 g, 2 hour criteria. Sensitivity ranged from 86 to 97 percent; specificity ranged from 79 to 87 percent. Prevalence ranged from 1.6 to 4.1 percent. PPV ranged from 7 to 20 percent; NPV ranged from 99 to 100 percent.

Three studies assessed a 50 g OGCT (≥140 mg/dL) with GDM confirmed using the WHO 75 g criteria. Sensitivity was 43 to 85 percent and specificity was 73 to 94 percent. Prevalence ranged from 3.7 to 15.7. In two studies with prevalence less than 10 percent, PPV was 18 and 20 percent; in one study in which prevalence was 10 or more, PPV was 58 percent. The median NPV for all studies was 99 percent.

Seven studies assessed FPG to screen for GDM; GDM was confirmed using CC criteria. Four FPG thresholds were compared— ≥85 mg/dL: sensitivity was 87 percent (95% CI, 81 to 91) and specificity was 52 percent (95% CI, 50 to 55); ≥90 mg/dL: sensitivity was 77 percent (95% CI, 66 to 85) and specificity was 76 percent (95% CI, 75 to77); ≥92 mg/dL: sensitivity was 76 percent (95% CI, 55 to 91) and specificity 92 percent (95% CI, 86 to 96); ≥95 mg/dL:

sensitivity was 54 percent (95% CI, 32 to 74) and specificity was 93 percent (95% CI, 90 to 96). While the effect on health outcomes was not part of this Key Question, the Toronto Trihospital and HAPO studies demonstrated the ability of using fasting glucose to predict GDM outcomes.

Limited data support the use of HbA1c as a screening test. One study conducted in the United Arab Emirates using an HbA1c value of 5.5 percent or more lacked specificity (21 percent) despite good sensitivity (82 percent). A study conducted in Turkey showed that an HbA1c cutoff of 7.2 percent or more had 64 percent sensitivity and specificity. HbA1c does not perform as well as the 50 g OGCT as a screening test for GDM. However, when HbA1c is markedly elevated, this supports a possible diagnosis of overt diabetes discovered in pregnancy. Since 2011–2012, the ADA has endorsed the use of an HbA1c of 6.5 percent or more as diagnostic of diabetes in nonpregnant women.[36]

Although eight studies examined risk factors for screening women, our review did not identify compelling evidence for or against risk factor-based screening. Studies used different diagnostic criteria and could not be pooled. Sensitivity and specificity varied widely across studies.

Only three studies included women who were in their first trimester of pregnancy, and they used different diagnostic criteria. Therefore, no conclusions can be made about the test characteristics of the screening tests for this group of women.

Four studies compared the 75 g and 100 g load tests, but they were conducted in different countries and used different criteria or thresholds. The prevalence of GDM ranged from 1.4 to 50 percent. Sensitivity and specificity varied widely across studies. Limited data are available to draw conclusions about the effectiveness of the different options for diagnostic testing for GDM. However, because both the 75 g and 100 g load tests are positively linked with outcomes[3,51] and the 75 g test is less time consuming, the adoption of the 75 g glucose load may be warranted, even if thresholds continue to be debated.[3,51]

The IADPSG has proposed the elimination of a screening test in favor of proceeding directly to a diagnostic test for GDM. We identified only one study that compared the IADPSG criteria with the Australasian Diabetes in Pregnancy Society (two-step) criteria. The sensitivity was 82 percent (95% CI: 74 to 88) and specificity was 94 percent (95% CI: 93 to 96); the PPV and NPV were 61 percent (95% CI: 53 to 68) and 98 (95% CI: 97 to 99), respectively.

Prevalence and Predictive Values

The prevalence of GDM varied across studies and the diagnostic criteria used. Factors contributing to the variability included differences in study setting (i.e., country), screening practices (e.g., universal vs. selective), and population characteristics (e.g., race/ethnicity, age, BMI).

The predictive value of a screening or diagnostic test is determined by the test's sensitivity and specificity and by the prevalence of GDM. Table B presents a series of scenarios that demonstrate the changes in PPV and NPV for three levels of prevalence (7 percent, 15 percent, and 25 percent).[6] Separate tables are presented for different screening and diagnostic criteria. The higher the prevalence of GDM, the higher the PPV, or the more likely a positive result is able to predict the presence of GDM. When the prevalence of GDM is low, the PPV is also low, even when the test has high sensitivity and specificity. Generally the NPV (negative result rules out GDM) is very high—98 percent or better at a GDM prevalence of 7 percent.

Table B. Relationship between predictive values and prevalence for different screening tests

Screening Test	Prevalence	Positive Predictive Value	Negative Predictive Value
50 g OGCT ≥140 mg/dL by CC/ADA (2000-2010) Sensitivity=85%; Specificity=86%	7%	31%	99%
	15%	52%	97%
	25%	67%	95%
50 g OGCT ≥130 mg/dL by CC/ADA (2000-2010) Sensitivity=99%; Specificity=77%	7%	24%	100%
	15%	43%	100%
	25%	59%	100%
50 g OGCT ≥140 mg/dL by NDDG Sensitivity=85%; Specificity=83%	7%	27%	99%
	15%	47%	97%
	25%	63%	94%
50 g OGCT ≥130 mg/dL by NDDG Sensitivity=88%; Specificity=66% (median)	7%	16%	99%
	15%	31%	97%
	25%	46%	94%
50 g OGCT ≥140 mg/dL by ADA 75 g Sensitivity=88%; Specificity=84% (median)	7%	29%	99%
	15%	49%	98%
	25%	65%	95%
50 g OGCT ≥140 mg/dL by WHO Sensitivity=78%; Specificity=81% (median)	7%	24%	98%
	15%	42%	95%
	25%	58%	92%
FPG (≥85 mg/dL) by CC/ADA (2000-2010) Sensitivity = 87%; Specificity = 52%	7%	12%	98%
	15%	24%	96%
	25%	38%	92%
Risk factor screening by various criteria Sensitivity=84%; Specificity=72% (median)	7%	21%	98%
	15%	38%	96%
	25%	54%	93%

ADA = American Diabetes Association; CC = Carpenter-Coustan; FPG = fasting plasma glucose; NDDG = National Diabetes Data Group; OGCT = oral glucose challenge test; WHO =World Health Organization

Key Question 2

Only two retrospective cohort studies were relevant to Key Question 2, which asked about the direct benefits and harms of screening for GDM. One retrospective cohort study (n=1,000) conducted in Thailand showed a significantly greater incidence of cesarean deliveries in the screened group. A survey of a subset of participants (n=93) in a large prospective cohort study involving 116,678 nurses age 25–42 years in the United States found the incidence of macrosomia (infant weight ≥ 4.3 kg) was the same in the screened and unscreened groups (7 percent each group).

No RCTs were available to answer questions about screening. There is a paucity of evidence on the effect of screening women for GDM on health outcomes. The comparison for this question was women who had and had not undergone screening. Since screening is now

commonplace it may be unlikely to identify studies or cohorts in which this comparison is feasible.

Key Question 3

Thirty-eight studies provided data for Key Question 3, which sought to examine health outcomes for women who met various criteria for GDM and did not receive treatment. A summary of the results is provided in Table D. The majority of data came from cohort studies or the untreated groups from RCTs. Study quality was assessed using the Newcastle-Ottawa Scale with a possible total of nine stars. The median quality score was 9 out of 9 stars. Studies receiving lower scores most often did not control for potential confounding, and/or had an important proportion of patients lost to followup. Overall, the majority of studies were considered good quality (36 of 38, 95 percent).

A wide variety of diagnostic criteria and thresholds were compared across the studies. The most common groups reported and compared were GDM diagnosed by CC criteria, no GDM by any criteria (normal), impaired glucose tolerance defined as one abnormal glucose value, and false positive (positive OGCT, negative OGTT). Only single studies contributed data for many of the comparisons and outcomes; therefore, results that showed no statistically significant differences between groups cannot be interpreted as equivalence between groups, and they do not rule out potential differences.

Two studies did not group women according to criteria (as above) but examined glucose levels as a continuous outcome and their association with maternal and neonatal outcomes. Both studies were methodologically strong. A continuous positive association was found between maternal glucose and birthweight (both studies), as well as fetal hyperinsulinemia (one study only). There was some evidence of an association between glucose levels and primary cesarean section and neonatal hypoglycemia, although the associations were not consistently significant. No clear glucose thresholds were found that were predictive of poor outcomes. One of these studies also found significantly fewer cases of preeclampsia, cesarean section, shoulder dystocia and/or birth injury, clinical neonatal hypoglycemia, and hyperbilirubinemia for women with no GDM compared with those meeting IADPSG criteria.

For maternal outcomes among the studies that compared groups as described above, women without GDM and those testing false positive showed fewer cases of preeclampsia than those meeting CC criteria. No differences in preeclampsia were found for other comparisons, although evidence was based on few studies per comparison.

Fewer cases of cesarean section were found among women without GDM compared with women meeting criteria for CC GDM, CC 1 abnormal OGTT, CC false positives, NDDG false positives, NDDG 1 abnormal oral glucose tolerance test, WHO IGT, IADPSG impaired fasting glucose (IFG), and IADPSG impaired glucose tolerance (IGT) IFG. There were fewer cases of cesarean section among false positives compared with women meeting criteria for CC GDM. For 12 other comparisons, there were no differences in rates of cesarean delivery.

For maternal hypertension, significant differences were found for 8 of 16 comparisons; many comparisons were based on single studies. No GDM groups showed lower incidence of maternal hypertension when compared with CC GDM, CC 1 abnormal OGTT, IADPSG IFG, IADPSG IGT-2 (double-impaired glucose tolerance), and IADPSG IGT IFG. Other comparisons showing significant differences were CC GDM versus false positives (lower incidence for false positives), IADPSG IGT versus IGT IFG (lower incidence for IGT), and IADPSG IFG versus IGT IFG (lower incidence for IFG).

Based on single studies, no differences were observed for maternal birth trauma for three comparisons. For maternal weight gain (less weight gain considered beneficial), significant differences were found for 3 of 12 comparisons: IADPSG IGT versus no GDM (favored IGT), IADPSG IFG versus no GDM (favored IFG), IADPSG IGT-2 versus no GDM (favored IGT-2). All comparisons were based on single studies. For maternal mortality/morbidity, single studies contributed to three comparisons, and no differences were found except for fewer cases among patient groups with no GDM compared with IADPSG GDM. No studies provided data on long-term maternal outcomes, such as type 2 diabetes mellitus, obesity, and hypertension.

The most commonly reported outcome for the offspring was macrosomia >4,000 g. Six of 11 comparisons showed a significant difference: there were fewer cases in the group without GDM compared with CC GDM, CC 1 abnormal OGTT, NDDG GDM (unrecognized), NDDG false positives, and WHO IGT. Fewer cases were found for women with false-positive results compared with CC GDM. Data for macrosomia >4,500 g were available for four comparisons and showed significant differences in two comparisons: patient groups with no GDM had fewer cases compared with women with CC GDM and with unrecognized NDDG GDM.

For shoulder dystocia, significant differences were found for 7 of 17 comparisons; all but one comparison were based on single studies. Patient groups with no GDM showed lower incidence of shoulder dystocia when compared with CC GDM (5 studies), NDDG GDM (unrecognized), NDDG false positive, WHO IGT, IADPSG IFG, and IADPSG IGT IFG. The other significant difference showed lower incidence among the false-positive group compared with CC 1 abnormal OGTT.

For fetal birth trauma or injury, four studies compared CC GDM, NDDG GDM, and WHO IGT with patient groups without GDM. No differences were observed except for NDDG GDM, which favored the group with no GDM. Only one difference was found for neonatal hypoglycemia, with fewer cases among patient groups without GDM compared with those meeting CC criteria. There were 16 comparisons for hyperbilirubinemia; the majority were based on single studies. Three comparisons showed significant differences between groups: patient groups with no GDM had fewer cases compared with CC false positive, IADPSG IGT, and IADPSG IGT-2, respectively. No differences were found for fetal morbidity/mortality for any of eight comparisons, which may be attributable to small numbers of events within some comparisons. Moreover, comparisons were based on single studies.

Based on a single study, significant differences were found in prevalence of childhood obesity for CC GDM versus patients without GDM (lower prevalence for no GDM) and CC GDM versus false positives (lower prevalence for false positives). This was consistent for both childhood obesity >85th percentile as well as >95th percentile. However, this study was unable to control for maternal weight or BMI, which are established predictors of childhood obesity. No differences, based on the same single study, were found for the other four comparisons within >85th or >95th percentiles. No other studies provided data on long-term outcomes, including type 2 diabetes mellitus and transgenerational GDM.

In summary, different thresholds of glucose intolerance affect maternal and neonatal outcomes of varying clinical importance. While many studies have attempted to measure the association between various criteria for GDM and pregnancy outcomes in the absence of treatment, the ability of a study or pooled analysis to find a statistically significant difference in pregnancy outcomes appears more dependent on study design, in particular the size of the study or pooled analysis, rather than the criteria used for diagnosing GDM. This is not surprising given the strong support found for a continuous positive relationship between glucose and a variety of

pregnancy outcomes. The clinical significance of absolute differences in event rates requires consideration by decisionmakers even though statistical significance was reached at the strictest diagnostic glucose thresholds for some outcomes.

This question focused on outcomes for women who did not receive treatment for GDM. While women with untreated GDM have a variety of poorer outcomes than women without GDM, it cannot be assumed that treatment of GDM reverses all the short- and long-term poor outcomes observed in women with untreated GDM. Some of the reasons for the poorer outcomes in women that have untreated GDM may not be modifiable, such as the influences of genetic makeup. The strength of evidence was insufficient for most outcomes and comparisons in this question due to high risk of bias (observational studies), inconsistency across studies, and/or imprecise results. The strength of evidence was low for the following outcomes and comparisons: preeclampsia (CC GDM vs. no GDM, CC GDM vs. false positives), macrosomia >4,000 g (CC GDM vs. no GDM, CC GDM vs. false positives, CC GDM vs. 1 abnormal OGTT, CC false positives vs. no GDM, NDDG false positives vs. no GDM), macrosomia >4,500 g (CC GDM vs. no GDM), and shoulder dystocia (CC GDM vs. no GDM).

Key Question 4

Eleven studies provided data for Key Question 4 to assess the effects of treatment for GDM on health outcomes of mothers and offspring. All studies compared diet modification, glucose monitoring, and insulin as needed with standard care. The strength of evidence for key outcomes is summarized in Table C, and a summary of the results is provided in Table D.

Among the 11 included studies, 5 were RCTs and 6 were cohort studies. The risk of bias for the RCTs was low for one trial, unclear for three trials, and high for one trial. The trials that were unclear most commonly did not report detailed methods for sequence generation and allocation concealment. The trial assessed as high risk of bias was due to lack of blinding for outcome assessment and incomplete outcome data. The six cohort studies were all considered high quality, with overall scores of 7 to 9 on a 9-point scale.

There was moderate evidence showing a significant difference for preeclampsia, with fewer cases in the treated group. There was inconsistency across studies in terms of differences in maternal weight gain, and the strength of evidence was considered insufficient. There were no data on long-term outcomes among women, including type 2 diabetes mellitus, obesity, and hypertension.

In terms of infant outcomes, there was insufficient evidence for birth trauma. This was driven by lack of precision in the effect estimates and inconsistency across studies: there was no difference for RCTs, but a significant difference favoring treatment in the one cohort study. The incidence of shoulder dystocia was significantly lower in the treated groups, and this finding was consistent for the three RCTs and four cohort studies. Overall, the evidence for shoulder dystocia was considered moderate, showing a difference in favor of the treated group. For neonatal hypoglycemia, the strength of evidence was low, suggesting no difference between groups. Moderate evidence showed benefits of treatment in terms of macrosomia (>4,000 g).

Only one study provided data on long-term metabolic outcomes among the offspring at a 7- to 11-year followup. The strength of evidence was insufficient. For both outcomes—impaired glucose tolerance and type 2 diabetes mellitus—no differences were found between groups although the estimates were imprecise. No differences were observed in single studies that assessed BMI >95 (7- to 11-year followup) and BMI >85 percentile (5- to 7-year followup). Overall, pooled results showed no difference in BMI, and the strength of evidence was low.

In summary, there was moderate evidence showing differences in preeclampsia and shoulder dystocia, with fewer cases among women (and offspring) who were treated compared with those not receiving treatment. There was also moderate evidence showing significantly fewer cases of macrosomia (>4,000 g) among offspring of women who received treatment for GDM. The results were driven by the two largest RCTs, the Maternal Fetal Medicine Unit (MFMU)[25] and the Australian Carbohydrate Intolerance in Pregnancy Study (ACHOIS),[52] which had unclear and low risk of bias, respectively. There was little evidence showing differences between groups in other key maternal and infant outcomes. One potential explanation is that for the most part, the study populations included women whose glucose intolerance was less marked, as those whose glucose intolerance was more pronounced would not have been entered into a trial in which they may be assigned to a group receiving no treatment. For outcomes where results were inconsistent between studies, different study glucose threshold entry criteria did not explain the variation. For some outcomes, particularly the long-term outcomes, the strength of evidence was insufficient or low, suggesting that further research may change the results and increase our confidence in them. Moreover, for some outcomes events were rare, and the studies may not have had the power to detect clinically important differences between groups; therefore, findings of no significant difference should not be interpreted as equivalence between groups.

Table C. Strength of evidence for Key Question 4: maternal and infant outcomes

Outcome	# Studies (# Patients)	Overall Strength of Evidence	Comment
Preeclampsia	3 RCTs (2,014)	moderate (favors treatment)	The evidence provides moderate confidence that the estimate reflects the true effect in favor of the treatment group.
	1 cohort (258)	insufficient	
Maternal weight gain	4 RCTs (2,530)	insufficient	There is insufficient evidence to draw conclusions for this outcome due to inconsistency across studies and imprecise effect estimates.
	2 cohorts (515)	insufficient	
Birth injury	2 RCTs (1,230)	low (no difference)	There is insufficient evidence to make a conclusion for this outcome. There is a difference in findings for the RCTs and cohort studies; the number of events and participants across all studies does not allow for a conclusion.
	1 cohort (389)	insufficient	
Shoulder dystocia	3 RCTs (2,044)	moderate (favors treatment)	The evidence provides moderate confidence that the estimate reflects the true effect in favor of the treatment group.
	4 cohorts (3,054)	low (favors treatment)	
Neonatal hypoglycemia	4 RCTs (2,367)	low (no difference)	The evidence provides low confidence that there is no difference between groups.
	2 cohorts (2,054)	insufficient	
Macrosomia (>4,000 g)	5 RCTs (2,643)	moderate (favors treatment)	The evidence provides moderate confidence that the estimate reflects the true effect in favor of the treatment group.
	6 cohorts (3,426)	low (favors treatment)	

Table C. Strength of evidence for Key Question 4: maternal and infant outcomes (continued)

Outcome	# Studies (# Patients)	Overall Strength of Evidence	Comment
Long-term metabolic outcomes: impaired glucose tolerance	1 RCT (89)	insufficient	There is insufficient evidence to draw conclusions for this outcome.
Long-term metabolic outcomes: type 2 diabetes mellitus	1 RCT (89)	insufficient	There is insufficient evidence to draw conclusions for this outcome.
Long-term metabolic outcomes: BMI (assessed as >85th and >95th percentile)	2 RCTs (284)	low (no difference)	The evidence provides low confidence that there is no difference between groups.

BMI = body mass index; RCT = randomized controlled trial

Key Question 5

Five studies (four RCTs and one cohort study) provided data for Key Question 5 on the harms associated with treatment of GDM. Among the four RCTs, one had low and three had unclear risk of bias. The cohort study was high quality (7/9 points); the primary limitation was not controlling for potential confounders.

Four of the studies provided data on the incidence of infants that were small for gestational age and showed no significant difference between groups. This finding may have resulted from inadequate power to detect differences due to a small number of events; therefore, the finding of no significant difference should not be interpreted as equivalence between groups.

Four of the studies provided data on admission to the NICU and showed no significant differences overall. One study was an outlier because it showed a significant difference favoring the no treatment group. This difference may be attributable to site-specific policies and procedures or lack of blinding of investigators to treatment arms. Two studies reported on the number of prenatal visits and generally found significantly more visits between the treatment groups.

Two of the RCTs showed no significant difference overall in the rate of induction of labor, although there was important statistical heterogeneity between studies. One RCT showed significantly more inductions of labor in the treatment group,[52] while the other study did not.[25] Different study protocols may account for the heterogeneity of results between studies. In the first study that showed more inductions of labor in the treatment group, no recommendations were provided regarding obstetrical care. In the second study, antenatal surveillance was reserved for standard obstetrical indications. Based on the studies included in Key Question 4 (five RCTs and six cohort studies), there was no difference in rates of cesarean section between treatment and nontreatment groups.

A single study assessed depression and anxiety at 6 weeks after study entry and 3 months postpartum using the Spielberger State-Trait Anxiety Inventory and the Edinburgh Postnatal Depression Score, respectively. There was no significant difference in anxiety between the groups at either time point, although there were significantly lower rates of depression in the treatment group at 3 months postpartum. These results should be interpreted cautiously because the assessment of depression and anxiety was conducted in a subgroup of the study population.

There was no evidence for some of the outcomes stipulated in the protocol, including costs and resource allocation.

Findings in Relationship to What Is Already Known

This review provides evidence that treating GDM reduces some poor maternal and neonatal outcomes. The recent MFMU trial[25] published in 2009 reinforces the findings of the earlier ACHOIS trial that was published in 2005[52] and included in an earlier version of this review.[24] Both trials showed that treating GDM to targets of 5.3 or 5.5 mmol/L fasting and 6.7 or 7.0 mmol/L 2 hours postmeal reduced neonatal birthweight, large for gestational age, macrosomia, shoulder dystocia, and preeclampsia, without a reduction in neonatal hypoglycemia or hyperbilirubinemia/jaundice requiring phototherapy, or an increase in small for gestational age. In contrast to the ACHOIS trial, MFMU demonstrated a reduced cesarean section rate in the GDM treatment group. The failure of ACHOIS to find a lower cesarean section rate despite reduced neonatal birthweight and macrosomia may have been the result of differing obstetrical practices or the different populations studied (e.g., the inclusion of some women with more marked glucose intolerance in ACHOIS, as reflected by the increased prevalence of insulin use; more black and Hispanic women in the MFMU study). Differences may have also resulted due to study design: in the ACHOIS trial, participants did not receive specific recommendations regarding obstetrical care, thus treatment was left to the discretion of the delivering health care provider. In the MFMU study, antenatal surveillance was reserved for standard obstetrical indications. Our findings of the effect of treatment of GDM is similar to a systematic review and meta-analysis published in 2010 by Horvath and colleagues.[53] This review included two older RCTs of GDM that were not included in our analysis because we restricted our inclusion criteria to studies published after 1995.

The HAPO Study Cooperative Research Group[3] used a simpler 75 g OGTT in a large international sample of women and confirmed findings of the earlier Toronto Trihospital study[51] that there is a continuous positive association between maternal glucose and increased birthweight, as well as fetal hyperinsulinemia (HAPO only), at levels below diagnostic thresholds for GDM that existed at the time of the study. However, no clear glucose thresholds were found for fetal overgrowth or a variety of other maternal and neonatal outcomes. Subsequently, the IADPSG developed diagnostic thresholds for GDM based on a consensus of expert opinion of what was considered to be the most important outcomes and the degree of acceptable risk for these outcomes. The thresholds chosen by the IADPSG were derived from the HAPO data to identify women with a higher risk (adjusted odds ratio 1.75) of large for gestational age, elevated c-peptide, and high neonatal body fat compared with the mean maternal glucose values of the HAPO study. The glucose threshold chosen by the IADPSG represents differing levels of risk for other outcomes. Specifically, their thresholds represent a 1.4 (1.26–1.56) risk for pregnancy-induced hypertension and a 1.3 (1.07–1.58) risk for shoulder dystocia. A dichotomous view of GDM may no longer be appropriate, given evidence of a continuous relationship between maternal blood glucose and pregnancy outcomes. An alternative approach may be to define different glucose thresholds based on maternal risk for poor pregnancy outcomes. This approach has been used in the context of lipid levels and risk of adverse cardiovascular outcomes.

Neither recent RCT was designed to determine diagnostic thresholds for GDM or therapeutic glucose targets. However, it is noteworthy that therapeutic glucose targets for both ACHOIS and MFMU were above the proposed diagnostic criteria of the IADPSG (fasting 5.5 mmol/L [99 mg/dL] and 5.3 mmol/L [95 mg/dL] and 2 hour postmeal of 7.0 mmol/L [126 mg/dL and 6.7 mmol/L 120 mg/dL], respectively). A change in diagnostic criteria without addressing management thresholds could contribute to clinical confusion. If diagnostic thresholds for GDM

below the treatment targets of the large RCTs are endorsed, this could ethically obstruct the possibility of future RCTs to compare different treatment targets above such diagnostic thresholds.

It has been hypothesized that treatment of GDM may reduce future poor metabolic outcomes for children born to mothers with GDM. If true, the potential for long-term gain is important from a clinical and public health perspective and may justify the "costs" of screening and treating women for GDM. However, the followup of offspring from two RCTs[52,54] and a HAPO cohort in Belfast [55] currently fail to support this hypothesis. This may be explained in part due to insufficient length of followup or inadequate numbers of events.

The HAPO study showed that maternal weight and glucose predict large for gestational age. However, BMI was the better predictor of large for gestational age than glucose until glucose thresholds higher than the diagnostic thresholds set by the IADPSG were reached.[56,57] Most cases of large for gestational age occur in neonates of mothers with normal glycemia. A large observational study found that the upper quartile of maternal BMI accounted for 23 percent of macrosomia, while GDM was responsible for only 3.8 percent.[58]

The ongoing obesity epidemic in the United States warrants careful consideration of a diagnostic approach for GDM that incorporates maternal BMI. This would require the development and validation of a risk model that incorporates maternal BMI as well as other modifiable risk factors. Such a model could facilitate the identification of women at high risk of adverse pregnancy outcomes and minimize exposure of lower risk women to unnecessary interventions.

Applicability

Several issues may limit the applicability of the evidence presented in this review to the U.S. population. All of the Key Questions asked about the effects of screening and treatment before and after 24 weeks' gestation. The vast majority of included studies screened women after 24 weeks' gestation; therefore, the results are not applicable to screening and treatment earlier in gestation.

For Key Question 1 on the test properties of screening and diagnostic tests, comparisons involving the WHO criteria are less applicable to the U.S. setting because these criteria are not used in North America. There were insufficient data from the included studies to assess the performance of screening or diagnostic tests for specific patient characteristics (e.g., BMI, race/ethnicity). Therefore it is unclear whether the evidence applies to specific subpopulations of women.

For Key Question 2, limited evidence was identified because the comparison of interest was women who had not undergone screening. Because screening is routine in prenatal care in the United States, the evidence (or limited evidence) is likely not helpful for U.S. decisionmaking, and a refinement of this question may be appropriate to reflect current practices and outstanding questions.

With respect to Key Question 3, all studies or groups included for analysis involved women who had not received treatment for GDM. It cannot be assumed that the same associations and outcomes would be observed in clinical practice in which standard care is to screen for and treat GDM. The untreated women may differ from the general population in ways that are related to the reasons for which they did not seek or receive early prenatal care (e.g., socioeconomic status). That is, the reasons they did not receive treatment for GDM are varied; some reasons, such as late presentation for obstetrical care, may confound the observed association with health

outcomes. Attempts were made to control for these factors in some studies (e.g., Langer and colleagues[59]) by including a group of women without GDM with similar known confounders or by adjusting for known confounders in the analysis. The adjusted estimates did not change the overall pooled results in the majority of cases and did not change the overall conclusions.

The majority of the studies for Key Questions 4 and 5 pertaining to the benefits and harms of treatment for GDM were conducted in North America or Australia. Most of the North American studies were inclusive of mixed racial populations and are likely applicable to the general U.S. population. Even though the Australian RCT[52] population had more white women with a lower BMI than the U.S. RCT (MFMU[25]), this should not affect applicability of most of their findings because these patient characteristics would be factors associated with lower risk of poor outcomes. Differences in physician or hospital billing structures between the United States and Australia may have accounted for the discrepant findings with respect to NICU admissions and, as a result, may limit the applicability of this finding in the United States. Among the studies included in Key Questions 4 and 5, a variety of glucose threshold criteria were used for inclusion, varying from 50 g screen positive with nondiagnostic OGTTs, to women who met NDDG criteria for a diagnosis of GDM. The two large RCTs[25,52] used different glucose thresholds for entry in their trials: WHO and CC criteria with a fasting glucose <95 mg/dL (5.3 mmol/L), respectively. The mean glucose levels at study entry were similar between these two RCTs, which may reflect a reluctance to assign women with more marked glucose intolerance to a group receiving no treatment. The results may not be applicable to women with higher levels of glucose intolerance.

Limitations of the Evidence Base

There is sparse evidence to clarify issues regarding the timing of screening and treatment for GDM (i.e., before and after 24 weeks' gestation). Earlier screening will help identify overt type 2 diabetes mellitus and distinguish this from GDM. This has important implications for clinical management and ongoing followup beyond pregnancy. Previously unrecognized type 2 diabetes mellitus diagnosed in pregnancy should be excluded from the diagnosis of GDM because this condition has the highest perinatal mortality rate of all classes of glucose intolerance in pregnancy.[60] This distinction within research studies will provide more targeted evidence to help obstetrical care providers to risk stratify obstetrical care and glycemic management of patients with overt type 2 diabetes mellitus diagnosed in pregnancy and those with less pronounced pregnancy-induced glucose intolerance. This will also facilitate better comparability across future studies. Few data were available on long-term outcomes. Furthermore, the studies included in this review do not provide evidence of a direct link between short-term and long-term outcomes (e.g., macrosomia and childhood obesity).

Care provider knowledge of the glucose screening and diagnostic results may have introduced a bias if their subsequent treatment of women differed depending on the results. This was of particular concern for Key Question 3, which assessed how the various criteria for GDM influenced pregnancy outcomes. For Key Question 3, many of the statistically significant differences seemed to be driven by the size of the study or pooled analysis (i.e., statistically significant differences could be found if the sample were sufficiently large). However, these differences may not be clinically important. The absolute differences in event rates between different glucose thresholds need careful consideration by decisionmakers, even though statistically significant differences were found. Another key limitation with the evidence for Key Question 3 is that the studies included were cohort studies, many of which did not control for

potential confounders. Therefore, any associations between glucose thresholds and outcomes should be interpreted with caution.

Given that the large landmark studies[51,61] show a continuous relationship between glucose and maternal and neonatal outcomes, the lack of clear thresholds contributes to the uncertainty regarding a diagnostic threshold for GDM. While there is controversy about where to set lower limits for diagnostic criteria, the identification of overt diabetes in pregnancy is imperative if this diagnosis has not occurred before pregnancy. Overt diabetes first identified in pregnancy should be distinguished from GDM to gain a better understanding of the true risk of GDM to pregnancy outcomes. Unfortunately there is no literature to guide diagnostic criteria for a diagnosis of overt diabetes in pregnancy.

There were several methodological concerns for this evidence base. For example, risk of spectrum bias and partial verification bias (Key Question 1); different definitions or methods of assessing key outcomes (e.g., clinical vs. biochemical neonatal hypoglycemia and hyperbilirubinemia) (Key Questions 3 and 4); and lack of blinding of treatment arms in some studies (Key Questions 4 and 5).

Future Research

Several important gaps in the current literature exist:

- The adoption of a consistent comparator for diagnosis of GDM, such as the 75 g OGTT, would facilitate comparisons across studies even if different diagnostic thresholds are used.
- Further analysis of the HAPO data could help answer some outstanding questions. For example, further analysis could better define absolute differences in rare event rates. This evidence could be used to inform discussions about the clinical importance of absolute differences in event rates at thresholds other than those of the IADPSG. Such analyses should include adjustment for important confounders such as maternal BMI.
- Further analysis of the HAPO data, examining center-to-center differences in glucose outcome relationships would be helpful in determining the usefulness of FPG as a screening test for GDM.
- Research is needed to clarify issues regarding earlier screening and treatment, particularly as they relate to the diagnosis, treatment, and long-term outcomes of pregestational (overt) diabetes.
- Further research of FPG, a screening test, is needed, given that the reproducibility of fasting glucose measurement is superior to postglucose load measurements.[62]
- Further study of the long-term metabolic outcomes in offspring whose mothers have been treated for GDM is warranted. In addition, data on the influences of GDM treatment on long-term breastfeeding success have not been studied. The association of breastfeeding with reduced poor metabolic outcomes in offspring of GDM has been found to have a dose-dependent response with duration of breastfeeding.[63]
- Implementation of well-conducted prospective cohort studies of the “real world” effects of GDM treatment on use of care is needed.
- Research on outcomes is needed to help determine the glucose thresholds and treatment targets at which GDM treatment benefits outweigh the risks of treatment and no treatment. This will best be achieved through well-conducted, large RCTs that randomize women with GDM to different glucose treatment targets.

- While this review did not identify evidence of substantial harms to treatment, the populations considered were mostly women whose GDM was controlled without medication. There is a risk for more precautionary management of women diagnosed with GDM, who are perceived by clinicians to be at greater risk, such as those managed with insulin, which may result in unnecessary interventions (e.g., cesarean section).[64] Therefore, RCTs investigating the care of women diagnosed with GDM, including fetal surveillance protocols, are needed to guide obstetrical investigations and management of GDM. Further, RCTs comparing delivery management for GDM with and without insulin or medical management are needed to provide clinicians guidance on appropriate timing and management of delivery in women with GDM to avoid unnecessary intervention in "the real world" driven by health care provider apprehension.
- The development of long-term studies that evaluate the potential increased or decreased resource use associated with the implementation of diabetes prevention strategies after a diagnosis of GDM is required.
- Studies to assess the long-term results that a label of GDM may have for future pregnancy planning, future pregnancy management, and future insurability are required.
- The increased prevalence of type 2 diabetes mellitus in women of reproductive age merits consideration of preconception screening for overt diabetes in women at risk of type 2 diabetes. In addition to poor maternal and neonatal outcomes associated with overt diabetes in pregnancy, there is potential for benefit of preconception care.
- Long-term benefits and harms need to be evaluated among different treatment modalities for GDM (e.g., diet, exercise, insulin, oral glucose-lowering medications, and/or combinations of these).
- Since 2011–2012, the American Diabetes Association has endorsed the use of an HbA1c of 6.5 percent or more as a diagnostic of diabetes in nonpregnant women.[36] Studies of HbA1c with trimester-specific cutoffs to determine the value at which overt diabetes should be diagnosed in pregnancy are needed.

Limitations of the Review

This review followed rigorous methodological standards, which were detailed a priori. The limitations of the review to fully answer the Key Questions are largely due to the nature and limitations of the existing evidence.

Several limitations need to be discussed regarding systematic reviews in general. First, there is a possibility of publication bias. The effects of publication bias on the results of diagnostic test accuracy reviews (Key Question 1) is not well understood, and the tools to investigate publication bias in these reviews have not been developed. For the remaining Key Questions, we may be missing unpublished and/or negative therapy studies and may be overestimating the benefits of certain approaches. However, we conducted a comprehensive and systematic search of the published literature for potentially relevant studies. Search strategies included combinations of subject headings and free text words. These searches were supplemented by handsearching for gray literature (i.e., unpublished or difficult-to-find studies). Despite these efforts, we recognize that we may have missed some studies.

There is also a possibility of study selection bias. However, we employed at least two independent reviewers and feel confident that the studies excluded from this report were done so for consistent and appropriate reasons. Our search was comprehensive, so it is unlikely that many studies in press or publication were missed.

Cost analysis of different screening and diagnostic approaches was not addressed in this review.

Conclusions

There was limited evidence regarding the test characteristics of current screening and diagnostic strategies for GDM. Lack of an agreed-upon gold standard for diagnosing GDM creates challenges for assessing the accuracy of tests and comparing across studies. The 50 g OGCT with a glucose threshold of 130 mg/dL versus 140 mg/dL improves sensitivity and reduces specificity (10 studies). Both thresholds have high negative predictive value, but variable positive predictive value across a range of GDM prevalence. There was limited evidence for the screening of GDM diagnosed less than 24 weeks' gestation (3 studies). Single studies compared the diagnostic characteristics of different pairs of diagnostic criteria in the same population. The use of fasting glucose (≥85 mg/dL) as a screen for GDM may be a practical alternative because of similar test characteristics to the OGCT, particularly in women who cannot tolerate any form of oral glucose load.

Evidence supports benefits of treating GDM, with little evidence of short-term harm. Specifically, treatment of GDM results in lower incidence of preeclampsia, macrosomia, and large for gestational age infants. Current research does not demonstrate a treatment effect of GDM on clinical neonatal hypoglycemia or future poor metabolic outcomes of the offspring. RCTs of GDM treatment show limited harm related to treating GDM, other than an increased demand for services. There is a risk for more precautionary management of women diagnosed with GDM, who are perceived by clinicians to be at greater risk, such as those managed with insulin, which may result in unnecessary interventions (e.g., cesarean section); however, this review found limited data for these outcomes, and further research on the care of women diagnosed with GDM (e.g., fetal surveillance protocols) is warranted.

What remains less clear is what the lower limit diagnostic thresholds for GDM should be. Given the continuous association between glucose and a variety of outcomes, decisions should be made in light of what outcomes altered by treatment are the most important and what level of increased risk is acceptable. A dichotomous view of GDM may no longer be appropriate, given evidence of a continuous relationship between maternal blood glucose and pregnancy outcomes. An alternative approach would be to define different glucose thresholds based on maternal risk for poor pregnancy outcomes.

Further study is needed regarding the long-term metabolic outcomes on offspring of mothers receiving GDM treatment; the "real world" impact of GDM treatment on use of care outside of structured research trials; and the results of the timing of screening for GDM, particularly before 24 weeks' gestation and in the first trimester of pregnancy. Early screening could help identify pregestational (i.e., overt) diabetes. Research is urgently required to determine the best way to diagnose and manage overt diabetes in pregnancy, particularly in an era of increasing rates of obesity and diabetes in the U.S. population.

Table D. Summary of evidence for all Key Questions

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ1. What are the sensitivities, specificities, reliabilities, and yields of current screening tests for GDM? (a) After 24 weeks' gestation? (b) During the first trimester and up to 24 weeks' gestation?	**(a) *After 24 wk gestation*** 51 prospective studies *Fair to good quality*	Limitations: Lack of an agreed upon gold standard for diagnosis of GDM creates challenges for assessing the accuracy of tests and comparing across studies. GDM was confirmed using criteria developed by CC, ADA, NDDG, and WHO. There were sparse data comparing overall approaches for diagnosis and screening, e.g., one-step vs. two-step, selective vs. universal. Consistency: Across studies numerous tests and thresholds were examined. Screening tests included the 50 g OGCT, FPG, risk factor-based screening, and other less common tests such as HbA1c, serum fructosamine.	Prevalence of GDM varied across studies and diagnostic criteria used. Results need to be interpreted in the context of prevalence. Comparisons involving WHO criteria are less applicable to the North American setting because these criteria are not used in North America.	• Prevalence varied across studies and diagnostic criteria: ADA 2000-2010 (75 g) 2.0 to 19% (range), CC 3.6 to 38%, NDDG 1.4 to 50%, WHO 2 to 24.5%. • 9 studies examined a 50 g OGCT with a cutoff value of ≥140 mg/dL; GDM was confirmed using CC criteria. Results: sensitivity 85%, specificity 86%, prevalence 3.8 to 31.9%, PPV 18 to 27% (prevalence <10), PPV 32 to 83% (prevalence ≥10), NPV median 98%. • 6 studies examined a 50 g OGCT (≥130 mg/dL); GDM was confirmed using CC criteria. Results: sensitivity 99%, specificity 77%, prevalence 4.3 to 29.5%, PPV 11 to 31% (prevalence <10), PPV 31 to 62% (prevalence ≥10), NPV median 100%. • 1 study examined a 50 g OGCT (≥200 mg/dL); GDM was confirmed using CC criteria. Sensitivity, specificity, PPV, and NPV were all 100%. Prevalence was 6.4%. • 7 studies examined a 50 g OGCT (≥140 mg/dL); GDM was confirmed using NDDG criteria. Results: sensitivity 85%, specificity 83%, prevalence 1.4 to 45.8%, PPV 12 to 39% (prevalence <10), PPV 57% (prevalence ≥10), NPV median 99%. • 3 studies examined a 50 g OGCT (≥130 mg/dL); GDM was confirmed using NDDG criteria. Results: sensitivity 67 to 90% (range), specificity 47 to 84%; prevalence 16.7 to 35.3%, PPV 20 to 75%, NPV 86 to 95%. • 3 studies examined a 50 g OGCT (different thresholds); GDM was confirmed using ADA 2000-2010 (75 g) criteria. Prevalence was 1.6 to 4.1% (range). Results: sensitivity 86 to 97% (range), specificity 79 to 87%; PPV 7 to 20%, NPV 99 to 100%.

Table D. Summary of evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ1. What are the sensitivities, specificities, reliabilities, and yields of current screening tests for GDM? (a) After 24 weeks' gestation? (b) During the first trimester and up to 24 weeks' gestation? (continued)	**(a) *After 24 wk gestation*** 51 prospective studies *Fair to good quality* (continued)	**(a) *After 24 wk gestation*** 51 prospective studies *Fair to good quality*		• 3 studies examined a 50 g OGCT (≥140 mg/dL); GDM was confirmed using WHO criteria. Results: sensitivity 43 to 85%, specificity 73 to 94%, prevalence 3.7 to 15.7%, PPV 18 to 20% (prevalence <10), PPV 58% (prevalence ≥10), NPV median 99%. • 7 studies examined FPG at different thresholds; GDM was confirmed using CC criteria. Results: at ≥85 mg/dL sensitivity 87%, specificity 52%; at ≥90 mg/dL sensitivity 77%, specificity 76%; at ≥92 mg/dL sensitivity 76%, specificity 92%; at ≥95 mg/dL sensitivity 54%, specificity 93%. At ≥85 mg/dL prevalence 1.4 to 34.53 (range). PPV 10% (prevalence <10) and 23 to 59% (prevalence ≥10). Median NPV 93%. • 8 studies examined risk factor-based screening but were not pooled. Studies used different criteria to confirm GDM. Results: sensitivity 48 to 95% (range), specificity 22 to 94%, prevalence 1.7 to 16.9%, PPV 5 to 19% (prevalence <10), PPV 20% (prevalence ≥10), NPV median 99%. • 1 study compared IADPSG vs. ADIPS 2 step (reference) to diagnose GDM. Results: sensitivity 82%, specificity 94%, prevalence 13.0%, PPV 61%, NPV 98%. • 4 studies compared 75 g and 100 g load tests to diagnose GDM. Prevalence ranged from 1.4 to 50%. Results were not pooled: sensitivity 18 to 100%, specificity 86 to 100%, PPV 12 to 100%, NPV 62 to 100%.

Table D. Summary of evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ1. What are the sensitivities, specificities, reliabilities, and yields of current screening tests for GDM? (a) After 24 weeks' gestation? (b) During the first trimester and up to 24 weeks' gestation? (continued)	***(b) During the first trimester and up to 24 wk gestation*** 3 prospective cohort studies	Limitations: Only 3 studies of women before 24 wks gestation; therefore, no conclusions can be made for test characteristics in early pregnancy. Consistency: Not applicable (not enough studies addressing the same question to judge consistency).	Evidence too limited to judge applicability.	• 1 study examined the 50 g OGCT at 10 wks and confirmed GDM using JSOG criteria (75 g). Results: sensitivity 88%, specificity 79%, prevalence 1.6%, PPV 7%, NPV 100%. • 1 study examined 50 g OGCT at 20 wks and confirmed GDM using ADA (2000-2010) 100 g criteria. Results: sensitivity 56%, specificity 94%, prevalence 3.6%, PPV 24%, NPV 98%. • 1 study compared 1st and 2nd trimester results using 3 screening tests (OGCT at ≥130 mg/dL, FPG, HbA1c); GDM confirmed using JSOG criteria. Results (OGCT) 1st trimester: prevalence 1.9%, sensitivity 93%, specificity 77%, PPV 7.1, NPV 99%; 2nd trimester: prevalence 2.9%, sensitivity 100%, specificity 85%, PPV 17%, NPV 100%.
KQ2: What is the direct evidence on the benefits and harms of screening women (before and after 24 weeks' gestation) for GDM to reduce maternal, fetal, and infant morbidity and mortality?	2 retrospective cohort studies *Fair and good quality*	Limitations: No RCTs available to answer this question. Consistency: Not applicable (not enough studies addressing the same question to judge consistency).	The comparison for this question was women who had and had not undergone screening. Since screening is now commonplace, it may be unlikely to identify studies or cohorts where this comparison is feasible.	1 study (n=1,000) showed more cesarean deliveries in the screened group. A second study (n=93) found the incidence of macrosomia (≥4.3 kg) was the same in screened and unscreened groups (7% each group). Based on the small number of studies and sample sizes, the effect of screening women for GDM on health outcomes is inconclusive.

Table D. Summary of evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ3: In the absence of treatment, how do health outcomes of mothers who meet various criteria for GDM and their offspring compare to those who do not meet the various criteria?	38 prospective or retrospective cohort studies; 2 studies were long-term followup from RCTs; however, only data from the untreated patients were included. *Fair to good quality*	Limitations: Strength of evidence was low to insufficient for all graded outcomes due to risk of bias (all observational studies), inconsistency, and/or imprecision. For many comparisons, the numbers of studies, participants, and/or events was low; therefore, findings of no statistically significant differences between groups do not imply equivalence or rule out potential differences. Consistency: A wide variety of diagnostic criteria and thresholds were compared across studies. There were often few studies with similar comparison groups. Differences in defining and assessing outcomes may have contributed to heterogeneity in results across studies (e.g., biochemical vs. clinical assessment of neonatal hypoglycemia).	All studies or groups included for analysis involved women who had not received treatment for GDM. These women may differ from the general population in other ways that are related to the reasons why they did not seek or receive early prenatal care (e.g., socioeconomic status).	*Maternal outcomes:* • A methodologically strong study showed a continuous positive relationship between increasing glucose levels and the incidence of primary cesarean section. This study also found significantly fewer cases of preeclampsia and cesarean section for women with no GDM vs. IADPSG. • For preeclampsia, significant differences were found for CC vs. patients with no GDM (3 studies), with fewer cases among the patients with no GDM, and for CC vs. false-positive groups (2 studies), with fewer cases among the false positives. The strength of evidence was low. No differences were found for NDDG false positive (2 studies), NDDG 1 abnormal OGTT vs. no GDM (1 study), or IGT WHO vs. no GDM (3 studies); the strength of evidence was insufficient. • For maternal weight gain, significant differences were found for 3 of 12 comparisons: IADPSG IGT vs. no GDM (favored IGT), IADPSG IFG vs. no GDM (favored IFG), IADPSG IGT-2 vs. no GDM (favored IGT-2). All comparisons were based on single studies (strength of evidence insufficient). *Fetal/neonatal/child outcomes:* • 2 methodologically strong studies showed a continuous positive relationship between increasing glucose levels and the incidence of macrosomia. 1 of these studies also showed significantly fewer cases of shoulder dystocia and/or birth injury, clinical neonatal hypoglycemia, and hyperbilirubinemia for women with no GDM vs. IADPSG.

Table D. Summary of evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ3: In the absence of treatment, how do health outcomes of mothers who meet various criteria for GDM and their offspring compare to those who do not meet the various criteria? (continued)	38 prospective or retrospective cohort studies; 2 studies were long-term followup from RCTs; however, only data from the untreated patients were included. *Fair to good quality* (continued)			• For macrosomia >4,000 g, 6 of 11 comparisons showed a significant difference: patient groups with no GDM had fewer cases compared with CC GDM (10 studies), CC 1 abnormal OGTT (7 studies), NDDG GDM (unrecognized) (1 study), NDDG false positives (4 studies), and WHO IGT (1 study). Fewer cases were found for women with false-positive results compared with CC GDM (5 studies). Data for macrosomia >4,500 g were available for 4 comparisons and showed significant differences in 2 cases: patient groups with no GDM had fewer cases compared with CC GDM (3 studies) and unrecognized NDDG GDM (1 study). The strength of evidence for macrosomia was low to insufficient. • For shoulder dystocia, significant differences were found for 7 of 17 comparisons; all comparisons but 1 were based on single studies (insufficient strength of evidence). Patient groups with no GDM showed lower incidence of shoulder dystocia when compared with CC GDM (5 studies, low strength of evidence), NDDG GDM (unrecognized), NDDG false positive, WHO IGT, IADPSG IFG, and IADPSG IGT IFG. The other significant difference showed lower incidence among the false-positive group compared with CC 1 abnormal OGTT.

Table D. Summary of evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ3: In the absence of treatment, how do health outcomes of mothers who meet various criteria for GDM and their offspring compare to those who do not meet the various criteria? (continued)	38 prospective or retrospective cohort studies; 2 studies were long-term followup from RCTs; however, only data from the untreated patients were included. *Fair to good quality* (continued)			• For fetal birth trauma/injury, single studies compared CC GDM and WHO IGT with no GDM and showed no differences. Two studies showed fewer cases for no GDM compared with NDDG GDM. Strength of evidence was insufficient for all comparisons. • No differences were found for neonatal hypoglycemia for any comparison, including CC GDM vs. no GDM (3 studies), CC GDM vs. 1 abnormal OGTT (1 study), CC 1 abnormal OGTT vs. no GDM (4 studies), NDDG GDM vs. no GDM (1 study), NDDG false positive vs. no GDM (1 study), and WHO IGT vs. no GDM (3 studies). Strength of evidence was insufficient for all comparisons.
KQ4: Does treatment modify the health outcomes of mothers who meet various criteria for GDM and offspring?	5 RCTs and 6 retrospective cohort studies. *Poor to good quality*	Limitations: For some outcomes, particularly the long-term outcomes, the strength of evidence was insufficient or low. Moreover, for some outcomes events were rare, and the studies may not have had the power to detect clinically important differences between groups; therefore, findings of no significant difference should not be interpreted as equivalence between groups.	For the most part, study populations included women whose glucose intolerance was less marked, as those whose glucose intolerance was more pronounced would not be entered into a trial in which they may be assigned to a group receiving no treatment. The majority of studies were conducted in North America or Australia, with 2 from Italy. Most of the North American studies were inclusive of mixed racial populations and are likely applicable to the general U.S. population.	*Maternal outcomes:* • Moderate evidence from 3 RCTs showed a significant difference for preeclampsia, with fewer cases in the treated group. • There was inconsistency across studies in terms of maternal weight gain (4 RCTs and 2 cohort studies); the strength of evidence was insufficient due to inconsistency and imprecision in effect estimates. *Offspring outcomes:* • There was insufficient evidence to make a conclusion for birth injury. There was inconsistency across studies, with the 2 RCTs showing no difference and the 1 cohort study showing a difference in favor of the treated group. The low number of events and participants across all studies resulted in imprecise estimates. • Moderate evidence showed significantly lower incidence of shoulder dystocia in the treated groups, and this finding was consistent for the 3 RCTs and 4 cohort studies.

Table D. Summary of evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ4: Does treatment modify the health outcomes of mothers who meet various criteria for GDM and offspring? (continued)	5 RCTs and 6 retrospective cohort studies. *Poor to good quality* (continued)	Consistency: Some inconsistency occurred at 2 levels. First, there were inconsistencies for some outcomes between RCTs and observational studies, which may be attributable to confounding and methods of selecting study groups (e.g., historical control groups). Second, in some instances there were inconsistencies across studies within designs, that were often attributable to the manner in which outcomes were defined or assessed (e.g., clinical vs. biochemical assessment of neonatal hypoglycemia).	Even though the Australian RCT population had more white women with a lower BMI than the U.S. RCTs; this should not affect applicability of most of their findings for the U.S. women because these subject characteristics would be factors associated with lower risk of poor outcomes.	• There was low evidence of no difference between groups for neonatal hypoglycemia based on 4 RCTs and 2 cohort studies. • For outcomes related to birthweight (including macrosomia >4,000 g, macrosomia >4,500 g, actual birthweight, and large for gestational age), differences were often observed favoring the treated groups. Strength of evidence was moderate for macrosomia >4,000 g. • 1 RCT followed patients for 7 to 11 years and found no differences for impaired glucose tolerance or type 2 DM, although the strength of evidence was considered insufficient. • No differences were observed in single studies that assessed BMI >95 (7-11 yr followup) and BMI >85 percentile (5-7 yr followup). Overall, pooled results showed no difference in BMI, and the strength of evidence was considered low.

Table D. Summary of evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ5: What are the harms of treating GDM and do they vary by diagnostic approach?	4 RCTs and 1 retrospective cohort study. *Fair to good quality*	*Limitations:* No study evaluated costs and resource allocation. Limited evidence on harms. Limited evidence for number of prenatal visits and NICU admissions. Findings of no significant differences may be attributable to low power and should not be interpreted as equivalence. *Consistency:* Not applicable (not enough studies addressing the same question to judge).	As above for KQ4. In addition, differences in billing structures between the United States and Australia may have accounted for the discrepant findings with respect to NICU admissions between these studies and as a result limit the applicability of this finding in the United States.	• 1 RCT assessed depression and anxiety at 6 weeks after study entry and 3 months postpartum. • There was no significant difference between groups in anxiety at either time point, although there were significantly lower rates of depression in the treatment group at 3 months postpartum. • 4 RCTs reported small for gestational age and found no significant difference. • 3 RCTs and 1 cohort study provided data on admission to NICU and showed no significant differences overall. One trial was an outlier because it showed a significant difference favoring the no treatment group. This difference may be attributable to site-specific policies and procedures. • 2 RCTs reported on the number of prenatal visits and generally found more visits among the treatment groups. • 2 RCTs reporting on induction of labor showed different results, with 1 showing a significant difference with more cases in the treatment group and the other showing no difference. • Based on studies included in KQ4, no differences between groups were found for cesarean section (5 RCTs, 6 cohorts) or unplanned cesarean section (1 RCT, 1 cohort).

ADA = American Diabetes Association; ADIPS = Australasian Diabetes in Pregnancy Society; BMI = body mass index; CC = Carpenter and Coustan; DM = type 2 diabetes mellitus; FPG = fasting plasma glucose; GDM = gestational diabetes mellitus; HbA1c = glycated hemoglobin; IADPSG = International Association of Diabetes in Pregnancy Study Groups; IFG = impaired fasting glucose; IGT = impaired glucose tolerance; IGT-2 = double impaired glucose tolerance; JSOG = Japan Society of Obstetrics and Gynecology; NDDG = National Diabetes Data Group; NPV = negative predictive value; NICU = neonatal intensive care unit; OGCT = oral glucose challenge test; OGTT = oral glucose tolerance test; PPV = positive predictive value; RCT = randomized controlled trial; wk(s) = week(s); WHO = World Health Organization

References

1. Balsells M, Garcia-Patterson A, Gich I, et al. Maternal and fetal outcome in women with type 2 versus type 1 diabetes mellitus: a systematic review and metaanalysis. J Clin Endocrinol Metab. 2009;94(11):4284-91. PMID: 19808847.
2. National Diabetes Data Group. Diabetes in America, 2nd ed. Bethesda, MD: National Institutes of Health; 1995.
3. HAPO Study Cooperative Research Group, Metzger B, Lowe L, et al. Hyperglycemia and adverse pregnancy outcomes. N Engl J Med. 2008;358(19):1991-2002. PMID: 18463375.
4. American Diabetes Association. Position statement: standards of medical care in diabetes - 2012. Diabetes Care. 2012;35(Suppl 1):S11-S63. PMID: 22187469.
5. Carpenter MW, Coustan DR. Criteria for screening tests for gestational diabetes. Am J Obstet Gynecol. 1982;144(7):768-73. PMID: 7148898.
6. Sacks DA, Hadden DR, Maresh M, et al. Frequency of gestational diabetes mellitus at collaborating centers based on IADPSG consensus panel-recommended criteria: the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study. Diabetes Care. 2012;35(3):526-8. PMID: 22355019.
7. Ferrara A. Increasing prevalence of gestational diabetes mellitus: a public health perspective. Diabetes Care. 2007;30(Suppl 2):S141-S146. PMID: 17596462.
8. Gillman MW, Oakey H, Baghurst PA, et al. Effect of treatment of gestational diabetes mellitus on obesity in the next generation. Diabetes Care. 2010;33(5):964-8. PMID: 20150300.
9. Kaufmann RC, Schleyhahn FT, Huffman DG, et al. Gestational diabetes diagnostic criteria: long-term maternal follow-up. Am J Obstet Gynecol. 1995;172(2 Pt 1):621-5. PMID: 7856695.
10. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2006;29(Suppl 1):S43-S48. PMID: 16373932.
11. Berger H, Crane J, Farine D, et al. Screening for gestational diabetes mellitus. J Obstet Gynaecol Can. 2002;24(11):894-912. PMID: 12417905.
12. Naylor CD, Sermer M, Chen E, et al. Selective screening for gestational diabetes mellitus. Toronto Trihospital Gestational Diabetes Project Investigators. N Engl J Med. 1997;337(22):1591-6. PMID: 9371855.
13. Hillier T, Vesco K, Pedula K, et al. Screening for gestational diabetes mellitus: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med. 2008;148:766-75. PMID: 18490689.
14. American College of Obstetricians and Gynecologists Committee on Practice Bulletins. ACOG Practice Bulletin. Clinical management guidelines for obstetrician-gynecologists. Number 30, September 2001. Gestational Diabetes. Obstet Gynecol. 2001;98(3):525-38. PMID: 1154779.
15. Meltzer SJ, Snyder J, Penrod JR, et al. Gestational diabetes mellitus screening and diagnosis: a prospective randomised controlled trial comparing costs of one-step and two-step methods. BJOG. 2010;117(4):407-15. PMID: 20105163.
16. Gabbe S, Gregory R, Power M, et al. Management of diabetes mellitus by obstetrician-gynecologists. Obstet Gynecol. 2004;103(6):1229-34. PMID: 15172857.
17. American Diabetes Association. Position Statement: Diabetes mellitus. Diabetes Care. 2004;27(Suppl 1):S11-S14. PMID: 14693922.
18. Moses RG, Cheung NW. Point: universal screening for gestational diabetes mellitus. Diabetes Care. 2009;32(7):1349-51. PMID: 19564479.
19. Danilenko-Dixon DR, Van Winter JT, Nelson RL, et al. Universal versus selective gestational diabetes screening: application of 1997 American Diabetes Association recommendations. Am J Obstet Gynecol. 1999;181(4):798-802. PMID: 10521732.

20. Berger H, Sermer M. Counterpoint: selective screening for gestational diabetes mellitus. Diabetes Care. 2009;32(7):1352-4. PMID: 19564480.

21. Metzger B, Gabbe S, Persson B, et al. International Association of Diabetes and Pregnancy Study Groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care. 2010;33(3):676-82. PMID: 20190296.

22. Lipscombe LL, Hux JE. Trends in diabetes prevalence, incidence, and mortality in Ontario, Canada 1995-2005: a population-based study. Lancet. 2007;369(9563):750-6. PMID: 17336651.

23. American Diabetes Association. Position statement: gestational diabetes mellitus. Diabetes Care. 2003;26(Suppl 1):S103-S105. PMID: 12502631.

24. U.S.Preventive Services Task Force. Screening for gestational diabetes mellitus: recommendations and rationale. Obstet Gynecol. 2003;101(393):395. PMID: 12576265.

25. Landon MB, Spong CY, Thom E, et al. A multicenter, randomized trial of treatment for mild gestational diabetes. N Engl J Med. 2009;361(14):1339-48. PMID: 19797280.

26. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 1999;22(Suppl 1):S5-S19.

27. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2000;23(Suppl 1):S4-S19. PMID: 12017675.

28. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2001;24(Suppl 1):S5-S20.

29. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2002;25(Suppl 1):S5-S20.

30. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2003;26(Suppl 1):S5-S20. PMID: 12502614.

31. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2004 Jan;27(Suppl 1):S5-S10. PMID: 14693921.

32. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2005;28(Suppl 1):S37-S42. PMID: 15618111.

33. American Diabetes Association. Diagnosis and classification of diabetes mellitus.Diabetes Care. 2007 Jan;30(Suppl 1):S42-S7. PMID: 17192378.

34. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2008;31(Suppl 1):S55-S60. PMID: 18165338.

35. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2009;32(Suppl 1):S62-S67. PMID: 19118289.

36. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2010;33(Suppl 1):S62-S69. PMID: 20042775.

37. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care. 2010;33(3):676-82. PMID: 20190296.

38. Jovanovic L. American Diabetes Association's Fourth International Workshop-Conference on Gestational Diabetes Mellitus: summary and discussion. Therapeutic interventions. Diabetes Care. 1998;21(Suppl 2):B131-37. PMID: 9704240.

39. Metzger BE, Oats JN, Kjos SL, et al. Summary and Recommendations of the Fifth International Workshop-Conference on Gestational Diabetes Mellitus. Diabetes Care. 2007;30(Suppl 2):S251-S260. PMID: 17596481.

40. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. Diabetes. 1979;28(12):1039-57. PMID: 510803.

41. World Health Organization. Definition, diagnosis and classification of diabetes mellitus and its complications. Report of a WHO Consultation. Part 1: Diagnosis and classification of diabetes mellitus. 1999.

42. World Health Organization. Report of a WHO study Group (Technical Report Series No.727). Report of a WHO study group (Technical Report Series No. 727). 1985.

43. Canadian Diabetes Association Clinical Practice Guidelines Expert Committee. Canadian Diabetes Association 2003 Clinical Practice Guidelines for the Prevention and Management of Diabetes in Canada. Can J Diabetes. 2003;27(Suppl 2), S1-S152.

44. Canadian Diabetes Association 2008 Clinical Practice Guidelines for the Prevention and Management of Diabetes in Canada [corrected] [published erratum appears in Can J Diabetes 2009 Mar;33(1):46]. Can J Diabetes. 2008;32:iv.

45. Sempowski IP, Houlden RL. Managing diabetes during pregnancy. Guide for family physicians. Canadian Family Physician Médecin De Famille Canadien. 2003;49:761-7. PMID: 12836864.

46. Metzger BE. Summary and recommendations of the Third International Workshop-Conference on Gestational Diabetes Mellitus. Diabetes. 1991;40(Suppl 2):197-201. PMID: 1748259.

47. Hoffman L, Nolan C, Wilson JD, et al. Gestational diabetes mellitus--management guidelines. The Australasian Diabetes in Pregnancy Society. The Medical Journal Of Australia. 1998;169(2):93-7. PMID: 9700346.

48. Brown CJ, Dawson A, Dodds R, et al. Report of the Pregnancy and Neonatal Care Group. Diabetic Medicine: A Journal Of The British Diabetic Association. 1996;13(9 Suppl 4):S43-S53. PMID: 8894455.

49. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36. PMID: 22007046.

50. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58(10):982-90. PMID: 16168343.

51. Sermer M, Naylor CD, Gare DJ, et al. Impact of increasing carbohydrate intolerance on maternal-fetal outcomes in 3637 women without gestational diabetes. The Toronto Tri-Hospital Gestational Diabetes Project. Am J Obstet Gynecol. 1995;173(1):146-56. PMID: 7631672.

52. Crowther CA, Hiller JE, Moss JR, et al. Effect of treatment of gestational diabetes mellitus on pregnancy outcomes. N Engl J Med. 2005;352(24):2477-86. PMID: 15951574.

53. Horvath K, Koch K, Jeitler K, et al. Effects of treatment in women with gestational diabetes mellitus: systematic review and meta-analysis. BMJ: British Medical Journal (International Edition). 2010;340:c1395. PMID: 20360215.

54. Malcolm JC, Lawson ML, Gaboury I, et al. Glucose tolerance of offspring of mother with gestational diabetes mellitus in a low-risk population. Diabetic Med. 2006;23(5):565-70. PMID: 16681566.

55. Pettitt DJ, McKenna S, McLaughlin C, et al. Maternal glucose at 28 weeks of gestation is not associated with obesity in 2-year-old offspring: The Belfast Hyperglycemia and Adverse Pregnancy Outcome (HAPO) family study. Diabetes Care. 2010;33(6):1219-23. PMID: 20215449.

56. Ryan EA. Diagnosing gestational diabetes. Diabetologia. 2011;54(3):480-6. PMID: 21203743.

57. HAPO Study Cooperative Research Group. Hyperglycaemia and Adverse Pregnancy Outcome (HAPO) Study: associations with maternal body mass index. BJOG. 2010;117(5):575-84. PMID: 20089115.

58. Ricart W, Lopez J, Mozas J, et al. Body mass index has a greater impact on pregnancy outcomes than gestational hyperglycaemia. Diabetologia. 2005;48(9):1736-42. PMID: 16052327.

59. Langer O, Yogev Y, Most O, et al. Gestational diabetes: the consequences of not treating. Am J Obstet Gynecol. 2005;192(4):989-97. PMID: 15846171.

60. Cundy T, Gamble G, Townend K, et al. Perinatal mortality in Type 2 diabetes mellitus. Diabet Med. 2000;17(1):33-9. PMID: 10691157.

61. Sacks DA, Greenspoon JS, bu-Fadil S, et al. Toward universal criteria for gestational diabetes: The 75-gram glucose tolerance test in pregnancy. Am J Obstet Gynecol. 1995;172(2 I):607-14. PMID: 7856693.

62. Rasmussen SS, Glumer C, Sandbaek A, et al. Short-term reproducibility of impaired fasting glycaemia, impaired glucose tolerance and diabetes The ADDITION study, DK. Diabetes Res Clin Pract. 2008;80(1):146-52. PMID: 18082284.

63. Schaefer-Graf UM, Hartmann R, Pawliczak J, et al. Association of breast-feeding and early childhood overweight in children from mothers with gestational diabetes mellitus. Diabetes Care. 2006;29(5):1105-7. PMID: 16644645.

64. Buchanan TA, Kjos SL, Montoro MN, et al. Use of fetal ultrasound to select metabolic therapy for pregnancies complicated by mild gestational diabetes. Diabetes Care. 1994;17(4):275-83. PMID: 8026282.

Introduction

Gestational Diabetes Mellitus

Gestational diabetes mellitus (GDM) is defined as glucose intolerance first discovered in pregnancy. Pregestational diabetes refers to any type of diabetes diagnosed before pregnancy. Pregnant women with pregestational diabetes experience an increased risk of poor maternal, fetal and neonatal outcomes.[1] The extent to which GDM predicts adverse outcomes for mother, fetus and neonate is less clear.

Depending on the diagnostic criteria used and the population screened, the prevalence of GDM ranges from 1.1 to 25.5 percent of pregnancies in the United States.[2-4] In 2009 the Centers for Disease Control and Prevention reported a prevalence of 4.8 percent of diabetes in pregnancy. An estimated 0.5 percent of these cases likely represented women with pregestational diabetes. Data from the international Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study[3] indicate that 6.7 percent of the women met a fasting plasma glucose threshold of 95 mg/dL (5.3 mmol/L), which is in keeping with the Carpenter and Coustan[5] (CC) criteria that are in common practice in North America. In contrast, 17.8 percent of women were diagnosed with GDM using the International Association of Diabetes in Pregnancy Study Groups (IADPSG) criteria in which lower glucose thresholds are proposed to diagnose GDM.

The prevalence of GDM is not only influenced by diagnostic criteria but also by population characteristics. In a recent publication, data from the HAPO study demonstrate wide variability in GDM prevalence across a variety of study centers internationally and within the United States, even when the same diagnostic criteria are applied (i.e., IADPSG).[6] Prevalence in the United States ranged from 15.5 percent in Providence, RI, to 25.5 percent in Bellflower, CA. There are ethnic differences in the prevalence of GDM in the United States. Native American, Asian, Hispanic, and African-American women are at higher risk than non-Hispanic white women based on CC criteria and/or hospital discharge diagnosis.[7] Data from 2000 showed that prevalence was highest among Asian and Hispanic women (~7 to 8 percent), intermediate among African-American women (~6 percent), and lower among non-Hispanic white women (~5 percent). The rate of increase of prevalence over the past 10 years has been highest for Asian and African-American women. A report from Montana demonstrated that the prevalence of GDM increased by approximately 10 percent among white women and by approximately 21 percent among Native American women from 2000 to 2003.[7]

The incidence of GDM has increased over the past decades in parallel with the increase in rates of obesity and type 2 diabetes mellitus, and this trend is expected to continue. In 2001 in the United States, the prevalence of obesity (body mass index [BMI] $\geq$30) was 20.9 percent and the prevalence of diabetes was 7.9 percent.[8] It is unclear how much the increase in obesity will impact the proportion of women diagnosed with overt diabetes during pregnancy versus transient pregnancy induced glucose intolerance.[9]

GDM is usually diagnosed after 20 weeks' gestation when placental hormones that have the opposite effect of insulin on glucose metabolism increase substantially. Women with adequate insulin secreting capacity overcome this insulin resistance of pregnancy by secreting more endogenous insulin in order to maintain normal blood glucose. Women with less adequate pancreatic reserve are unable to produce adequate insulin to overcome the increase in insulin resistance, and glucose intolerance results.

Glucose abnormalities in women with GDM usually resolve postpartum, but commonly recur in subsequent pregnancies. Women with GDM have an increased risk of future development of overt diabetes. The cumulative incidence of diabetes after a diagnosis of GDM varies widely depending on maternal BMI, ethnicity, and time since index pregnancy, and may reach levels as high as 60 percent.[10] When glucose abnormalities persist postpartum in a woman with GDM, her diabetes is recategorized as overt diabetes. When this occurs, the possibility that this woman had pregestational (i.e., overt) diabetes increases, especially if the diagnosis of GDM occurred prior to 20 weeks' gestation and glucose levels were markedly elevated in pregnancy.

The increased rates of obesity and type 2 diabetes mellitus, particularly among young females, makes it increasingly important to distinguish the effect of obesity and pregestational diabetes from GDM.[11,12] There is considerable variability in the proportion of women with suspected pregestational diabetes among studies that investigate pregnancy outcomes of women with GDM. This contributes to the confusion surrounding the true morbidity of GDM. In an attempt to enable better comparability across future studies and more accurate risk stratification of pregnant women with diabetes, recommendations[13] have proposed the exclusion of women with more severe glucose abnormalities in pregnancy from the diagnosis of GDM in an attempt to exclude women with pregestational (i.e., overt diabetes) from the population of women defined as having GDM. This proposal is in contrast to the older definition of GDM as any degree of glucose intolerance first discovered in pregnancy.

Risk Factors

Risk factors for GDM include greater maternal age, higher BMI, member of an ethnic group at increased risk for development of type 2 diabetes mellitus (i.e., Hispanic, African, Native American, South or East Asian, or Pacific Inlands ancestry), polyhydramnios, past history of GDM, macrosomia in a previous pregnancy, history of unexplained stillbirth, type 2 diabetes mellitus in a first degree relative, polycystic ovary syndrome, and metabolic syndrome.[14] Low risk of GDM is usually defined as young (age less than 25 or 30 years), non-Hispanic white, normal BMI (25 kg/m^2 or less), no history of previous glucose intolerance or adverse pregnancy outcomes associated with GDM, and no first degree relative with known diabetes.[7,15] Women at high risk of GDM are usually defined as having multiple risk factors for GDM. Women at moderate risk of GDM do not satisfy all criteria of women at low risk, but they lack two or more risks for GDM.

Screening and Diagnostic Strategies

The 2008 U.S. Preventive Services Task Force (USPSTF) evidence review on screening for GDM concluded that, at that time, "evidence was insufficient to assess the balance of benefits and harms of screening for gestational diabetes mellitus either before or after 24 weeks' gestation."[16] The report suggested that "…until there was better evidence clinicians should discuss screening for GDM with their patient and make case-by-case decisions. Discussions should include information about the uncertainty of benefits and harm as well as the frequency of positive screening test results."

The 2001 practice guidelines of the American College of Obstetricians and Gynecologists (ACOG) endorsed risk factor-based screening for GDM, recognizing that low risk women may be less likely to benefit from screening with glucose measurements. Women were considered low risk of GDM if they met all the following criteria: (1) younger than 25 years; (2) not a member of an ethnic group at high risk for development of type 2 diabetes mellitus; (3) BMI of

25 kg/m^2 or less; (4) no history of previous glucose intolerance or adverse pregnancy outcomes associated with GDM; and (5) no first degree relative with known diabetes. AGOG will update their 2001 practice guidelines on GDM based on the proceedings of the 2012 National Institutes of Health consensus conference on GDM diagnosis. Until 2011 the American Diabetes Association (ADA) also endorsed no screening for pregnant woman who met all the criteria mentioned above for low risk of GDM. In 2011 the ADA changed their recommendations to endorse glucose testing for GDM in all pregnant women who do not have a diagnosis of pregestational diabetes.

Common practices of glucose screening for GDM in North America involve a two-step approach in which patients with abnormal results on a screening test receive a subsequent diagnostic test.[17] Typically, a 50 g oral glucose challenge test (OGCT) is initially administered between 24 and 28 weeks' gestation in a nonfasting state, in women at moderate risk (i.e., women who do not meet all low risk criteria but lack two or more risk factors for GDM). The test is administered earlier in gestation for women at high risk of GDM (i.e., multiple risk factors for GDM) and repeated at 24-28 weeks' gestation if initial surveillance is normal. Patients who meet or exceed a screening threshold (usually 130 mg/dL or 140 mg/dL) receive a more involved diagnostic test, the oral glucose tolerance test (OGTT) in which a 75 g or 100 g oral glucose load is administered in a fasting state, and plasma glucose levels are evaluated after 1, 2, or 3 hours. A diagnosis of GDM is made in pregnant women when one or more glucose values fall at or above the specified glucose thresholds. Alternatively, a one-step method in which all patients or high risk patients forego the screening test and proceed directly to the OGTT has been recommended.[18] Interest has grown in assessing the usefulness of fasting plasma glucose as an alternative to the OGCT for screening for GDM for a number of reasons. First, the IADPG has proposed the use of a high threshold fasting plasma glucose 126 mg/dL (7.0 mmol/L) as soon as pregnancy is confirmed in women at high risk of type 2 diabetes mellitus as a means of identifying women with overt diabetes that likely predates their pregnancy. It is hypothesized that lesser degrees of fasting glucose elevation could be used to screen for GDM if this test is already being done to rule out overt diabetes. However, fasting glucose in early pregnancy is not well studied. Second, the reproducibility of fasting glucose measurement is superior to post glucose load measurements.[149] Third, some women do not tolerate the oral glucose drinks.

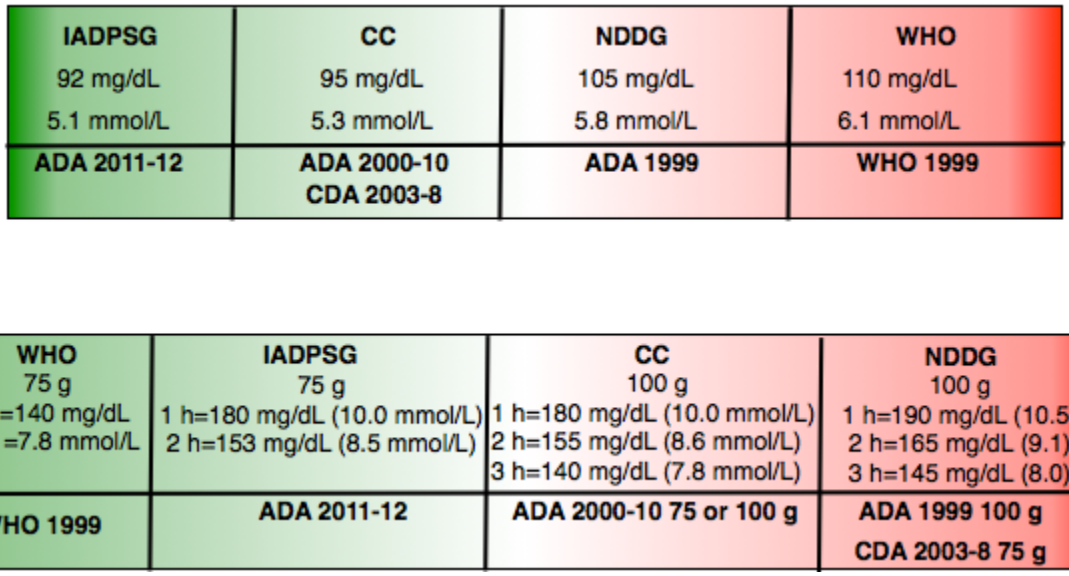
The absence of a universally accepted "gold standard" for the diagnosis of GDM has resulted in a variety of recommended diagnostic glucose thresholds that have been endorsed by different stakeholders (Table 1; Figure 1). These criteria reflect changes that have occurred in laboratory glucose measurements over the years, and new evidence that suggests the ability of different glucose thresholds to predict poor pregnancy outcomes. The different diagnostic criteria and thresholds result in different estimates of prevalence of GDM.

In 2004, a cross-sectional study reported that universal screening was the most common practice in the United States with 96 percent of obstetricians routinely screening for GDM.[19] In contrast, the guidelines of ACOG and the ADA at that time stated that women at low risk for GDM were unlikely to benefit from screening.[17,20] Since only 10 percent of pregnant women were categorized as low risk, some argued that selective screening contributed to confusion with little benefit and potential for harm.[21] Of particular concern was the association between risk factor-based screening and high rates of false negative results.[22] Others have endorsed alternative risk scoring systems for screening.[23]

The IADPSG, an international consensus group with representation from multiple obstetrical and diabetes organizations, recently spearheaded a re-examination of the definition of GDM in

an attempt to bring uniformity to GDM diagnoses.[24] The IADPSG recommended that a one-step 75 g OGTT be given to all pregnant women who do not have a diagnosis of overt diabetes. They also recommended that a single glucose value, rather than at least two abnormal values at or above diagnostic glucose thresholds on the OGTT be accepted as sufficient for a diagnosis of GDM. The diagnostic glucose thresholds recommended by the IADPSG were the maternal glucose values from the HAPO study[3] that identified a 1.75-fold increase (adjusted odds ratio relative to the mean cohort glucose values) in large for gestational age, elevated C-peptide, high neonatal body fat, or a combination of these factors. Since overt diabetes is often asymptomatic, may not have been screened for prior to conception, has a prevalence that is increasing dramatically in reproductive age women, and carries a higher risk for poor pregnancy outcomes, the IADPSG also recommended that all or at least women from high risk groups for type 2 diabetes mellitus be screened for overt diabetes at their first prenatal visit and excluded from the diagnosis of GDM using one of the following criteria: fasting plasma glucose ≥126 mg/dL (7.0 mmol/L), glycated hemoglobin (HbA1c) ≥6.5 percent (Diabetes Chronic Complications Trial/United Kingdom Prospective Diabetes Study standardized), or a random plasma glucose ≥200 mg/dL (11.1 mmol/L) confirmed by one of the first two measures.[25]

Figure 1. Comparison of different diagnostic thresholds for GDM



ADA = American Diabetes Association, CC = Carpenter-Coustan, CDA = Canadian Diabetes Association, dL= deciliter, g = grams, IADPSG = International Association of Diabetes in Pregnancy Study Groups, L= liter; mg = milligrams, mmol = millimoles; NDDG = National Diabetes Data Group, WHO = World Health Organization
Note: This figure presents the various diagnostic criteria for GDM. The top bar compares fasting glucose diagnostic thresholds. The bottom bar compares post glucose load diagnostic thresholds. The criteria are arranged from left (green) to right (red) from the lowest diagnostic glucose thresholds to the highest. The post glucose load bar is not entirely comparable because different glucose loads were used as indicated. The bottom part of each box shows which diagnostic thresholds were accepted by various organizations over the years including any modifications to the criteria. For example, ADA 2000 to 2010 endorsed the CC diagnostic thresholds on a 75g or 100g OGTT.

Table 1. Diagnostic criteria and plasma glucose thresholds for GDM

Organization	Year	Testing Schedule	Abnormal Value(s)	Threshold (Equal to or Greater Than)			
				0 (h)	1 (h)	2 (h)	3 (h)
ADA	1999[26]	50 g OGCT	1	—	140 mg/dL 7.8 mmol/L	—	—
		100 g OGTT	2 or more	105 mg/dL 5.8 mmol/L	190 mg/dL 10.5 mmol/L	165 mg/dL 9.1 mmol/L	145 mg/dL 8.0 mmol/L
ADA Low risk† excluded	2000-2010[13,27-36]	50 g OGCT	1	—	130 mg/dL 7.2 mmol/L or 140 mg/dL 7.8 mmol/L	—	—
		100 g or 75 g OGTT after overnight fast ≥8 hr	2 or more	95 mg/dL 5.3 mmol/L	180 mg/dL 10.0 mmol/L	155 mg/dL 8.6 mmol/L	140 mg/dL 7.8 mmol/L (3 hr value only for 100-g test)
IADPSG ADA	2011-2012[37]	75 g OGTT	1 or more	92 mg/dL 5.1 mmol/L	180 mg/dL 10.0 mmol/L	153 mg/dL 8.5 mmol/L	—
1. CC 2. 4th IWC (same) 3. 5th IWC (same as 4th but 75 g accepted with same glucose thresholds)	1. 1982[5] 2. 1998[38] 3. 2007[39]	50 g OGCT	1	—	130 mg/dL 7.2 mmol/L	—	—
		100 g OGTT	2 or more	95 mg/dL 5.3 mmol/L	180 mg/dL 10.0 mmol/L	155 mg/dL 8.6 mmol/L	140 mg/dL 7.8 mmol/L
NDDG	1979[40]	50 g OGCT	—	—	—	—	—
		100 g OGTT	2 or more	105 mg/dL 5.8 mmol/L	190 mg/dL 10.5 mmol/L	165 mg/dL 9.1 mmol/L	145 mg/dL 8.0 mmol/L
WHO	1999 WHO consultation[41]	75 g OGTT	1	6.1 mmol/L for IGT of pregnancy; 7.0 mmol/L for Dx of DM	—	140 mg/dL 7.8 mmol/L for IGT of pregnancy; 200 mg/dL 11.1 mmol/L for Dx of DM	—
WHO	1985 WHO study group report[42]	75 g OGTT	1	7.8 mmol/L 140 mg/dL for IGT of pregnancy	—	7.8 mmol/L (140 mg/dL); for IGT of pregnancy; 200 (11.1 mmol/L) for Dx of DM	—

Table 1. Diagnostic criteria and plasma glucose thresholds for GDM (continued)

Organization	Year	Testing Schedule	Abnormal Value(s)	Threshold (Equal to or Greater Than)			
				0 (h)	1 (h)	2 (h)	3 (h)
CDA	2003, 2008[43,44]	50 g OGCT	1	—	140 mg/dL 7.8 mmol/L or 186 mg/dL, 10.3 mmol/L Dx GDM	—	—
		75 g	2 or more	95 mg/dL 5.3 mmol/L	191 mg/dL 10.6 mmol/L	160 mg/dL 8.9 mmol/L	—
ACOG – risk factor 4th IWC	2001[17,45]	50 g	1	—	130 mg/dL 7.2 mmol/L or 140 mg/dL 7.8 mmol/L	—	—
		100 g CC	2 or more	95 mg/dL 5.3 mmol/L	180 mg/dL 10.0 mmol/L	155 mg/dL 8.5 mmol/L	140 mg/dL 7.8 mmol/L
		100 g NDDG	2 or more	105 mg/dL 5.8 mmol/L	190 mg/dL 10.5 mmol/L	165 mg/dL 9.1 mmol/L	145 mg/dL 8.0 mmol/L
3rd IWC	1991[46]	100 g OGTT	2 or more	105 mg/dL 5.8 mmol/L	190 mg/dL 10.5 mmol/L	165 mg/dL 9.1 mmol/L	145 mg/dL 8.0 mmol/L
ADIPS	1998[47]	50 g or 75 g nonfasting	1	—	140 mg/dL 7.8 mmol/L (50 g) or 144 mg/dL 8.0 mmol/L (75 g)	—	—
		75 g fasting	1	99 mg/dL 5.5 mmol/L	—	144 mg/dL 8.0 mmol/L or 1 62 mg/dL 9.0 mmol/L*	—

Table 1. Diagnostic criteria and plasma glucose thresholds for GDM (continued)

Organization	Year	Testing Schedule	Abnormal Value(s)	Threshold (Equal to or Greater Than)			
				0 (h)	1 (h)	2 (h)	3 (h)
EASD	1996[48]	75 g	1	108 mg/dL 6.0 mmol/L	—	162 mg/dL 9.0 mmol/L	—
USPSTF (Grade 1 recommendation)	2008‡	Risk Assessment 50 g OGCT	1	—	130 mg/dL 7.2 mmol/L or 140 mg/dL 7.8 mmol/L	—	—
		100 g OGTT	2 or more	NR	NR	NR	NR

ACOG = American College of Obstetricians and Gynecologists, ADA = American Diabetes Association, ADIPS = Australasian Diabetes in Pregnancy Society, CC = Carpenter, Coustan, CDA = Canadian Diabetes Association, DM = diabetes mellitus, Dx = diagnosis, EASD = European Association for the Study of Diabetes, h = hours; IADPSG = International Association of Diabetes in Pregnancy Study Groups, IGT = impaired glucose tolerance, IWC = International Workshop Conference, NDDG = National Diabetes Data Group, NR = not reported, OGCT = oral glucose challenge test, OGTT = oral glucose tolerance test, USPSTF = U.S. Preventive Services Task Force, WHO = World Health Organization

†Low risk defined as: (1) age <25 yr, (2) normal body weight, (3) no first degree relative with DM, (4) no history of abnormal glucose, (5) no history of poor obstetrical outcomes, (6) not of high-risk ethnicity for DM.

*In New Zealand.

‡ Screening for gestational diabetes mellitus: U.S. Preventive Services Task Force recommendation statement. Annals of Internal Medicine 2008;148(10):759-65.

Treatment Strategies

Initial treatment for GDM involves diet modification, glucose monitoring, and moderate exercise. When dietary management does not achieve desired glucose control, insulin or oral antidiabetic medications may be used.[49] Increased prenatal surveillance may also occur as well as changes in delivery management depending on fetal size and the effectiveness of measures to control glucose.

The 2008 USPSTF report found that treatment of women with mild GDM (excluding women who met World Health Organization criteria for overt diabetes) diagnosed after 24 weeks' gestation provided benefits in terms of maternal and neonatal health outcomes.[16] Specifically, they found evidence from a high quality trial involving 1,000 women showing a reduction in "any serious perinatal complication" which included death, shoulder dystocia, bone fracture, and nerve palsy.[50] The number of events for many of the individual outcomes was extremely small, which did not provide adequate evidence to make conclusions for individual outcomes. The same study showed a reduction in maternal hypertension.[50] Further, among a subset of survey respondents, mothers who received treatment were less depressed at 3 months and data showed a trend to better quality of life compared with women who did not receive treatment.[50]

The USPSTF report found no evidence of harms of treatment, although the available evidence was sparse and the review authors observed that these events may be rare and may not be observed in trials.[16] Potential harms of treatment may include small for gestational age neonates, maternal stress, and additional costs including those associated with laboratory testing as well as patient and clinician time.[51] Clinician time can include the physician as well as diabetes educators, nutritionists, and other providers of obstetrical care. Healthcare provider anxiety over the diagnosis of GDM is a potential harm that could result in additional, and possibly unnecessary or overly aggressive, fetal, and neonatal surveillance and delivery management. Evidence suggests that the label of GDM, regardless of need, appears to influence the care provided as evidenced by higher neonatal intensive care unit admission rates for the newborn babies of women treated for GDM.[52]

Scope and Key Questions

Scope of the Review

Based on systematic reviews published in 2003 and 2008, the USPSTF concluded that there was insufficient evidence upon which to make a recommendation regarding routine screening of all pregnant women for gestational diabetes.[16,53] However, several key studies have been published since the 2008 report.[3,9,54] The National Institutes of Health Office of Medical Applications of Research (OMAR) commissioned this report (Key Questions 3 to 5, see section below) and it was conducted by the Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Center (EPC) Program. OMAR will use the review to inform a consensus meeting and guideline development. The USPSTF joined this effort and will use the review to update its recommendation on screening for GDM (Key Questions 1 and 2 below).

The primary aims of this review were to: (1) identify the test properties of screening and diagnostic tests for GDM, (2) evaluate the potential benefits and harms of screening at ≥24 weeks and <24 weeks' gestation,(3) assess the impact of different screening and diagnostic thresholds on outcomes for mothers and their offspring, and (4) determine the effects of

treatment in modifying outcomes for women diagnosed with GDM. The benefits and harms of treatments will be considered in this review in order to determine the downstream effects of screening on health outcomes. The intent of this review was also to assess whether evidence gaps of the previous USPSTF reviews have been filled. These gaps included lack of sufficient evidence to determine whether maternal or fetal complications are reduced by screening; lack of screening studies with adequate power to evaluate health outcomes such as mortality, NICU admissions, hyperbilirubinemia; limited evidence on the accuracy of screening strategies; and insufficient evidence on the benefits of treating GDM in improving health outcomes.

Key Questions

The Key Questions for this evidence synthesis were developed by OMAR and the USPSTF to inform consensus meetings and guideline development (OMAR specifically developed Key Questions 3 to 5). Investigators from the University of Alberta EPC worked in consultation with representatives from AHRQ, OMAR and the USPSTF, and a panel of technical experts to operationalize the Key Questions. The technical expert panel provided content and methodological expertise throughout the development of this evidence synthesis. Participants of this panel are identified in the front matter of this report. The Key Questions are as follows:

Key Question 1: What are the sensitivities, specificities, reliabilities, and yields of current screening tests for GDM? (a) After 24 weeks' gestation? (b) During the first trimester and up to 24 weeks' gestation?

- Population: Pregnant women (≥24 weeks' gestation and <24 weeks' gestation) without known preexisting diabetes mellitus (DM)
- Interventions: Any screening or diagnostic test, including one-step, two-step, or other approach
- Comparators: Any reference standard
- Outcomes: Sensitivity, specificity, positive predictive value, negative predictive value, reliability (i.e., accuracy), and yield (i.e., prevalence)
- Timing: Any duration of followup
- Settings: All settings

Key Question 2: What is the direct evidence on the benefits and harms of screening women (before and after 24 weeks' gestation) for GDM to reduce maternal, fetal, and infant morbidity and mortality?

- Population: Pregnant women (≥24 weeks' gestation and <24 weeks' gestation) without known preexisting DM
- Interventions: Any screening or diagnostic test, including one-step, two-step, or other approach; if diagnosed with GDM, any treatment
- Comparators: No test for GDM
- Outcomes: Maternal, fetal, and infant morbidity and mortality
- Timing: Any duration of followup
- Settings: All settings

Key Question 3: In the absence of treatment, how do health outcomes of mothers who meet various criteria for GDM and their offspring compare to those who do not meet the various criteria?

- Population: Pregnant women (≥24 weeks' gestation and <24 weeks' gestation) without known preexisting DM who meet different test thresholds for GDM
- Interventions: None
- Comparators: Pregnant women (≥24 weeks' gestation and <24 weeks' gestation) without known preexisting DM who do *not* meet specific test thresholds for GDM
- Outcomes:
    - Maternal
        - Short-term: preeclampsia/maternal hypertension, cesarean delivery (elective and medically indicated), depression, birth trauma, mortality, weight gain
        - Long-term: type 2 DM risk, obesity, hypertension
    - Fetal/neonatal/child
        - Short-term: macrosomia, shoulder dystocia, clavicular fracture, brachial plexus injury (permanent and transient), birth injury, hypoglycemia, hyperbilirubinemia, mortality
        - Long-term: obesity, type 2 DM, transgenerational GDM
- Timing: Any duration of followup
- Settings: All settings

Key Question 4: Does treatment modify the health outcomes of mothers who meet various criteria for GDM and offspring?

- Population: Pregnant women (≥24 weeks' gestation and <24 weeks' gestation) without known preexisting DM who meet any diagnostic threshold for GDM
- Interventions: Any treatment for GDM including, but not limited to, dietary advice, blood glucose monitoring, insulin therapy, and oral hypoglycemic agents
- Comparators: Placebo or no treatment
- Outcomes:
    - Maternal
        - Short-term: preeclampsia/maternal hypertension, cesarean delivery (elective and medically indicated), depression, birth trauma, mortality, weight gain
        - Long-term: type 2 DM risk, obesity, hypertension
    - Fetal/neonatal/child
        - Short-term: macrosomia, shoulder dystocia, clavicular fracture, brachial plexus injury (permanent and transient), birth injury, hypoglycemia, hyperbilirubinemia, mortality
        - Long-term: obesity, type 2 DM, transgenerational GDM
- Timing: Any duration of followup
- Settings: All settings

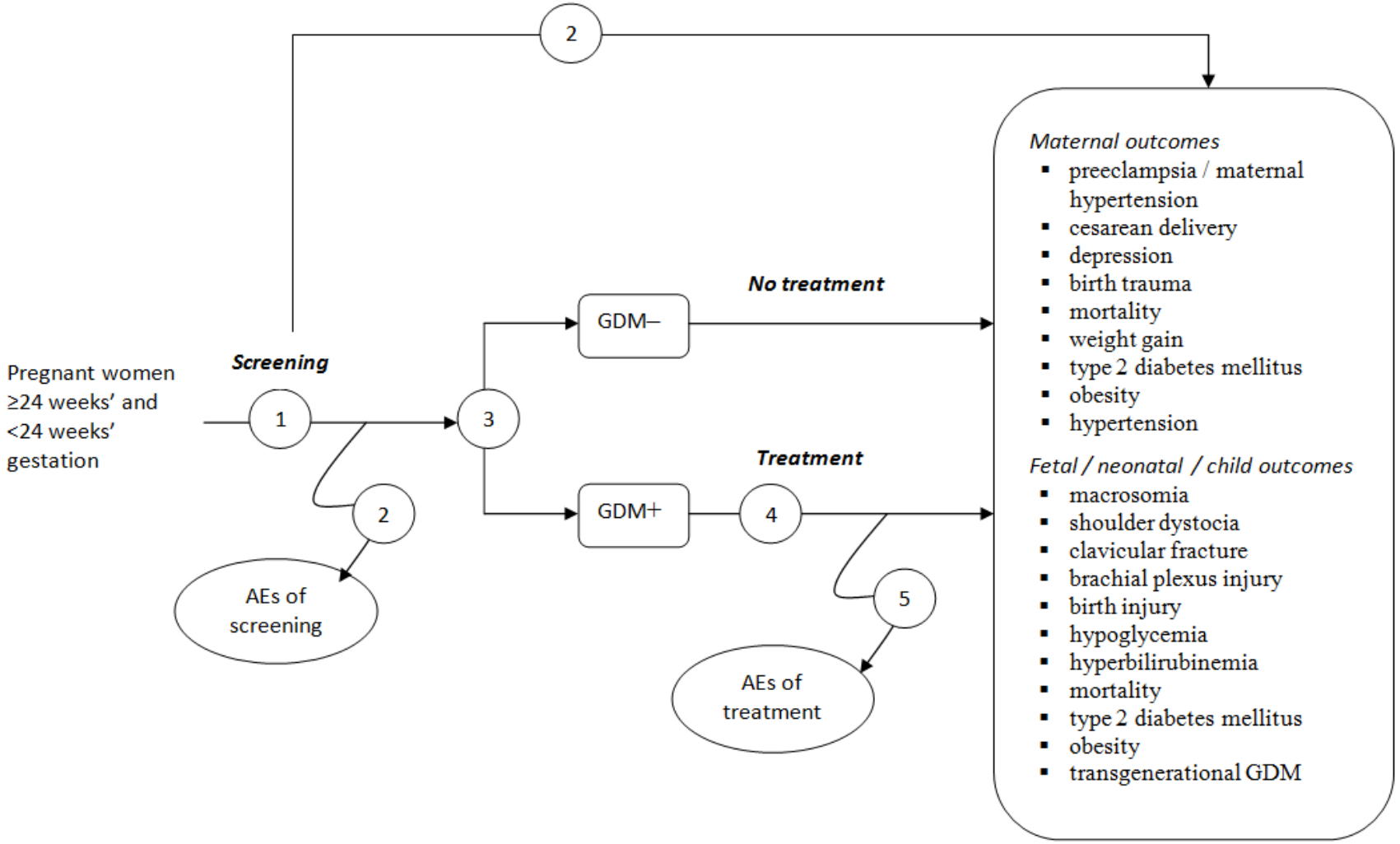
Key Question 5: What are the harms of treating GDM and do they vary by diagnostic approach?

- Population: Pregnant women (≥24 weeks' gestation and <24 weeks' gestation) without known preexisting DM who meet any diagnostic threshold for GDM
- Interventions: Any treatment for GDM including, but not limited to, dietary advice, blood glucose monitoring, insulin therapy, and oral hypoglycemic agents
- Comparators: Placebo or no treatment

- Outcomes: Harms, including anxiety, healthcare system issues, burden on practitioner's office, increased interventions due to treatment bias (e.g., increased cesarean sections resulting from bias of caregivers toward expectation of adverse outcomes), postpartum depression, SGA, costs, and resource allocations
- Timing: Any duration of followup
- Settings: All settings

We developed an analytic framework (Figure 2) to describe the path from screening pregnant women to the potential benefits and harms of treatment. The figure illustrates the clinical concepts and mechanism by which screening and treatment for GDM may result in beneficial or adverse maternal and fetal/neonatal/child outcomes. The figure also indicates the relation between the Key Questions and the specific links along the pathway from screening to final outcome.

Figure 2. Analytic framework for screening and diagnosing GDM



Note: The circled numbers correspond to the Key Questions.
AE = adverse event, GDM = gestational diabetes mellitus

Methods

The methods of this evidence synthesis are based on the methods outlined in the Agency for Healthcare Research and Quality (AHRQ) Methods Guide for Effectiveness and Comparative Effectiveness Reviews (www.effectivehealthcare.ahrq.gov/methodsguide.cfm) and the U.S. Preventive Services Task Force (USPSTF) Procedure Manual (www.uspreventiveservicestaskforce.org/uspstf08/methods/procmanual.pdf). The main sections in this chapter reflect the elements of the protocol established for the review. The methods and analyses were determined a priori, except where otherwise specified.

Topic Refinement and Technical Expert Panel

The National Institutes of Health Office of Medical Applications of Research (OMAR) commissioned this report and it was conducted by AHRQ through the Evidence-based Practice Center (EPC) Program. The Key Questions were developed by OMAR (Key Questions 3 to 5) and the USPSTF. OMAR will use the review to inform a consensus meeting and guideline development. The USPSTF joined this effort and will use the review to update its recommendation on screening for gestational diabetes mellitus.

Investigators from the University of Alberta EPC worked in consultation with representatives from AHRQ, OMAR and the USPSTF, and a panel of Technical Experts to operationalize the Key Questions. The Technical Expert Panel provided content and methodological expertise throughout the development of this evidence synthesis.

Literature Search Strategy

Our research librarian systematically searched the following bibliographic databases for studies published from 1995 to May 2012: MEDLINE® Ovid, Ovid MEDLINE® In-Process & Other Non-Indexed Citations, Cochrane Central Register of Controlled Trials (contains the Cochrane Pregnancy and Childbirth Group, which hand searches journals pertinent to its content area and adds relevant trials to the registry), Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts of Reviews of Effects (DARE), Global Health, Embase, Pascal CINAHL Plus with Full Text (EBSCO host), BIOSIS Previews® (Web of Knowledge℠), Science Citation Index Expanded® and Conference Proceedings Citation Index- Science (both via Web of Science℠), PubMed®, LILACS (Latin American and Caribbean Health Science Literature), National Library of Medicine (NLM) Gateway, and OCLC ProceedingsFirst and PapersFirst. We searched trial registries, including the WHO International Clinical Trials Registry Platform (ICTRP), ClinicalTrials.gov, and Current Controlled Trials.

We limited the search to trials and cohort studies published in English. For the search strategies, the research librarian developed a combination of subject headings and keywords for each electronic resource (see Appendix A for the detailed search strategies). The search strategies were not peer reviewed.

We searched the Web sites of relevant professional associations and research groups, including the American Diabetes Association, International Association of the Diabetes in Pregnancy Study Groups, International Symposium on Diabetes in Pregnancy, and Australasian Diabetes in Pregnancy Society for conference abstracts and proceedings from the past 3 years. We reviewed the reference lists of relevant reviews (including the 2008 USPSTF review) and included studies to identify additional studies.

We used Reference Manager® for Windows version 11.0 (2004–2005 Thomson ResearchSoft) bibliographic database to manage the results of our literature searches.

Inclusion and Exclusion Criteria

The research team developed the review eligibility criteria in consultation with the technical expert panel. The inclusion and exclusion criteria are presented in Table 2. We included studies only when less than 20 percent of enrolled women had a known history of pre-existing diabetes or separate data were provided for women with no pre-existing diabetes.

We limited our eligibility criteria to studies published in English due to lack of translation resources. This decision was made in consultation with the technical expert panel, which expressed no concerns that limiting the search to English language would forfeit important studies. We included studies that were published since 1995 in order to capture several key studies that were published in the late 1990s.

Randomized controlled trials (RCTs), nonrandomized controlled trials (NRCTs), and prospective and retrospective cohort studies were eligible for inclusion.

Table 2. Eligibility criteria for the review

Category	Criteria
Publication type	Primary research published in English from 1995 onward. Full text reports available (abstracts and conference proceedings excluded).
Study designs	RCTs, NRCTs, PCS, RCS.
Population	Pregnant women ≥24 weeks' gestation or <24 weeks' gestation, with no known history of pre-existing diabetes.
Comparators	**KQ1:** Any GDM screening or diagnostic test vs. any GDM reference standard or other screening or diagnostic test; **KQ2:** Any GDM screening test vs. no GDM screening test; **KQ3:** Women who meet various thresholds for GDM vs. those who do not meet various criteria for GDM, where women in both groups receive no treatment; **KQ4 and 5:** Any treatment for GDM, including but not limited to dietary advice, blood glucose monitoring, insulin therapy (all preparations), and oral hypoglycemic agents, vs. placebo or no treatment.
Outcomes	**KQ1:** Sensitivity, specificity, predictive values, accuracy, and yield (i.e., prevalence) **KQ2:** Maternal, fetal, and infant morbidity and mortality. **KQ3 and 4**: ***Maternal outcomes:*** *Short-term:* preeclampsia/maternal hypertension, cesarean delivery (elective and medically indicated), depression, birth trauma, mortality, weight gain; *Long-term:* type 2 DM risk, obesity, hypertension. ***Fetal, neonatal, and child:*** *Short-term:* macrosomia, shoulder dystocia, clavicular fracture, brachial plexus injury (permanent and transient), birth injury, hypoglycemia, hyperbilirubinemia, mortality; *Long-term:* obesity, type 2 DM, transgenerational GDM. **KQ5:** Harms, including anxiety, healthcare system issues, burden on practitioner's office, increased interventions due to treatment bias, postpartum depression, SGA, costs, and resource allocations.
Timing	Any duration of followup.
Setting	All settings are eligible.

DM = diabetes mellitus, GDM = gestational diabetes mellitus, KQ = Key Question, NRCT = nonrandomized controlled trials, PCS = prospective cohort study, RCS = retrospective cohort study, RCT = randomized controlled trial, SGA = small for gestational age

Study Selection

We assessed the eligibility of articles in two phases. In the first phase, two reviewers used broad criteria to independently screen the titles, keywords, and abstracts (when available) (Appendix B1). They rated each article as "include," "exclude," or "unclear." We retrieved the

full text article for any study that was classified as "include" or "unclear" by at least one reviewer. Two reviewers independently assessed each full text article using a detailed form (Appendix B2). We resolved disagreements by discussion and consensus or third-party adjudication.

Quality Assessment of Individual Studies

Two reviewers independently assessed the methodological quality of the studies and resolved discrepancies by discussion and consensus. We tested each quality assessment tool on a sample of studies and developed guidelines for assessing the remaining studies. In addition, we extracted the source of funding for each study. For studies included in Key Questions 2 to 5, we summarized the quality as "good," "fair," or "poor" based on assessments from the tools described below.

Quality Assessment of Diagnostic Studies

We assessed the methodological quality of studies relevant to Key Question 1 using the quality assessment of diagnostic accuracy studies (QUADAS)-2 checklist.[55] The tool consists of 14 items addressing important common biases in diagnostic studies such as spectrum, incorporation, verification, disease progression, and information biases. Individual items are rated "yes," "no," or "unclear" (Appendix B3a).

Quality Assessment of Trials

We assessed the internal validity of RCTs and NRCTs using the Cochrane Collaboration Risk of Bias tool (Appendix B3b). This tool consists of seven domains of potential bias (sequence generation, allocation concealment, blinding or participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting, and "other" sources of bias) and a categorization of the overall risk of bias.

Each domain was rated as having "low," "unclear," or "high" risk of bias. We assessed the blinding and incomplete outcome data items separately for subjective outcomes (e.g., depression scale) and objective clinical outcomes (e.g., mortality). We reported any additional sources of bias, such as baseline imbalances or design-specific risks of bias, in the "other" sources of bias domain.

The overall risk of bias assessment was based on the responses to individual domains. If one or more of the individual domains had a high risk of bias, we rated the overall score as high risk of bias. We rated the overall risk of bias as low only if all components were assessed as having a low risk of bias. The overall risk of bias was unclear in all other situations.

Quality Assessment of Cohort Studies

We used the Newcastle-Ottawa Quality Assessment Scale (Appendix B3c) to assess the methodological quality of prospective and retrospective cohort studies. The scale comprises eight items that evaluate three domains of quality: sample selection, comparability of cohorts, and assessment of outcomes. Each item that is adequately addressed is awarded one star, except for the "comparability of cohorts" item, for which a maximum of two stars can be given.

The overall score is calculated by tallying the stars. We considered a total score of 7 to 9 stars to indicate high quality, 4 or 6 stars to indicate moderate quality, and 3 or fewer stars to indicate poor quality.

Data Extraction

We extracted data using a structured, electronic form and imported the data into a Microsoft Excel™ 2007 spreadsheet (Microsoft Corp., Redmond, WA) (Appendix B4). One reviewer extracted data, and a second reviewer checked the data for accuracy and completeness. Reviewers resolved discrepancies by discussion and consensus or in consultation with a third party. We extracted the following data: author identification, year of publication, source of funding, study design, population (e.g., inclusion and exclusion criteria, number of patients enrolled, study withdrawals, duration of followup), patient baseline characteristics (e.g., age, race, ethnicity, weight, body mass index, previous diagnosis of gestational diabetes mellitus (GDM), family history of diabetes, comorbidities, smoking prevalence), details of the screening or diagnostic test and reference standard, glucose threshold for GDM, type of treatment, and outcomes, including adverse events.

We reported outcomes only if quantitative data were reported or could be derived from graphs. We did not include outcomes that were described only qualitatively (e.g., if study authors reported that "there was no difference between the groups") or for which only a p-value was reported.

We planned to extract any cost-related data, including costs to patients, insurance, or health care system, that were reported in the included studies. However, we did not search for cost effectiveness studies or conduct cost-effectiveness analyses of different treatment strategies. Studies that reported only costs and provided no other outcome data were not included in the review.

When more than one publication reported the results of a single study, we considered the earliest published report of the main outcome data to be the primary publication. We extracted data from the primary publication first and then any additional outcome data reported in the secondary publications.

Data Synthesis

We made the following assumptions and performed the following imputations to transform reported data into the form required for analysis. We extracted data from graphs using the measurement tool of Adobe Acrobat 9 Pro (Adobe Systems Inc., California, U.S.) when data were not reported in text or tables. As necessary, we approximated means by medians and used 95% confidence intervals (CI), p-values, or inter-quartile ranges to calculate or approximate standard deviations when they were not given. We calculated p-values when they are not reported.[56]

For Key Question 1, we constructed 2x2 tables and calculated sensitivity, specificity, positive and negative predictive values, accuracy (true positive plus true negative divided by the sum of true positive, true negative, false positive, and false negative) and yield (i.e., prevalence) of the screening or diagnostic tests. If studies were clinically homogenous, we pooled sensitivities and specificities using a hierarchical summary receiver-operator curve and bivariate analysis of sensitivity and specificity.[57]

We described the results of studies qualitatively and in evidence tables. For Key Questions 3to 5, we performed meta-analysis to synthesize the available data when studies were sufficiently similar in terms of their study design, population, screening or diagnostic test, and outcomes. This was done using the Mantel-Haenszel method for relative risks and the inverse variance

method for pooling mean differences. Due to the expected between-study differences, we decided a priori to combine results using the random effects model.[58]

We measured statistical heterogeneity among studies using the I^2 statistic. We considered an I^2 value of 75 percent or greater to represent substantial heterogeneity and did not pool studies indicating substantial heterogeneity. When studies were not pooled due to substantial heterogeneity, we performed subgroup analyses if the number of studies was sufficient to warrant these analyses.[59] Factors to be considered for subgroup analyses included glucose thresholds for tests, type of treatment, maternal age, race or ethnicity, and weight or body mass index, previous diagnosis of GDM, family history of diabetes, and comorbidities, which were extracted from each study.

We used Review Manager Version 5.0 (The Cochrane Collaboration, Copenhagen, Denmark) to perform meta-analyses. For dichotomous outcomes, we computed relative risks to estimate between-group differences. If no event was reported in one treatment arm, a correction factor of 0.5 was added to each cell of the 2x2 table in order to obtain estimates of the relative risk. For continuous variables, we calculated mean differences for individual studies. We reported all results with 95% CI.

Where possible, we assessed publication bias both visually using the funnel plot and quantitatively using Begg's[60] and Egger's[61] tests. Review Manager version 5.0.22 (The Cochrane Collaboration, Copenhagen, Denmark) and Stata version 7.0 (Stata Corp., College Station, TX) were used for all these analyses. In the event that studies could not be pooled, a narrative summary of the results was presented.

Strength of the Body of Evidence

Two independent reviewers graded the strength of evidence for major outcomes and comparisons for Key Questions 3 and 4 using the EPC GRADE (Grading of Recommendations Assessment, Development, and Evaluation) approach. We resolved discrepancies by discussion and consensus. We graded the evidence for the following key outcomes: birth injury, preeclampsia, neonatal hypoglycemia, maternal weight gain, and long-term metabolic outcomes of the child and mother. We made a post hoc decision to grade shoulder dystocia and macrosomia. These were not included in the protocol as outcomes that would be graded but were felt by the clinical investigators to be important to grade.

For each outcome, we assessed four major domains: risk of bias (rated as low, moderate, or high), consistency (rated as consistent, inconsistent, or unknown), directness (rated as direct or indirect), and precision (rated as precise or imprecise). No additional domains were used.

Based on the individual domains, we assigned the following overall evidence grades for each outcome for each comparison of interest: high, moderate, or low confidence that the evidence reflects the true effect. When no studies were available or an outcome or the evidence did not permit estimation of an effect, we rated the strength of evidence as insufficient.

To determine the overall strength of evidence score, we first considered the risk of bias domain. RCTs with a low risk of bias were initially considered to have a "high" strength of evidence, whereas RCTs with high risk of bias and well-conducted cohort studies received an initial grade of "moderate" strength of evidence. Low quality cohort studies received an initial grade of "low" strength of evidence. The strength of evidence was then upgraded or downgraded depending on the assessments of that body of evidence on the consistency, directness, and precision domains.

Applicability

We assessed the applicability of the body of evidence following the PICOTS (population, intervention, comparator, outcomes, timing of outcome measurement, and setting) format used to assess study characteristics. Factors that may potentially weaken the applicability of studies may include study population factors (e.g., race or ethnicity, age, risk level of GDM [i.e., weight, body mass index, previous GDM diagnosis, family history of diabetes], comorbidities), study design (i.e., highly controlled studies [e.g., RCTs] vs. observational studies), setting (e.g., primary vs. tertiary care), and experience of care providers.

Peer Review and Public Commentary

Peer reviewers were invited to provide written comments on the draft report based on their clinical, content, or methodologic expertise. Peer review comments on the draft report were addressed by the EPC in preparation of the final draft of the report. Peer reviewers did not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments are documented and will be published 3 months after the publication of the Evidence Report.

Potential reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Invited peer reviewers may not have any financial conflict of interest greater than $10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through AHRQ's public comment mechanism.

The draft report was posted for public commentary. Comments on the draft report were considered by the EPC in preparing the final report.

Results

This chapter reports on the results of our literature review and synthesis. First, we describe the results of our literature search and selection process. Description of the characteristics and methodological quality of the studies follow. We present our analysis of the study results by Key Question. Metagraphs and tables reporting the strength of evidence for key outcomes are available within each applicable section. Within each metagraph, the studies that provided data are indexed by the name of the first author. A list of abbreviations is provided at the end of the report.

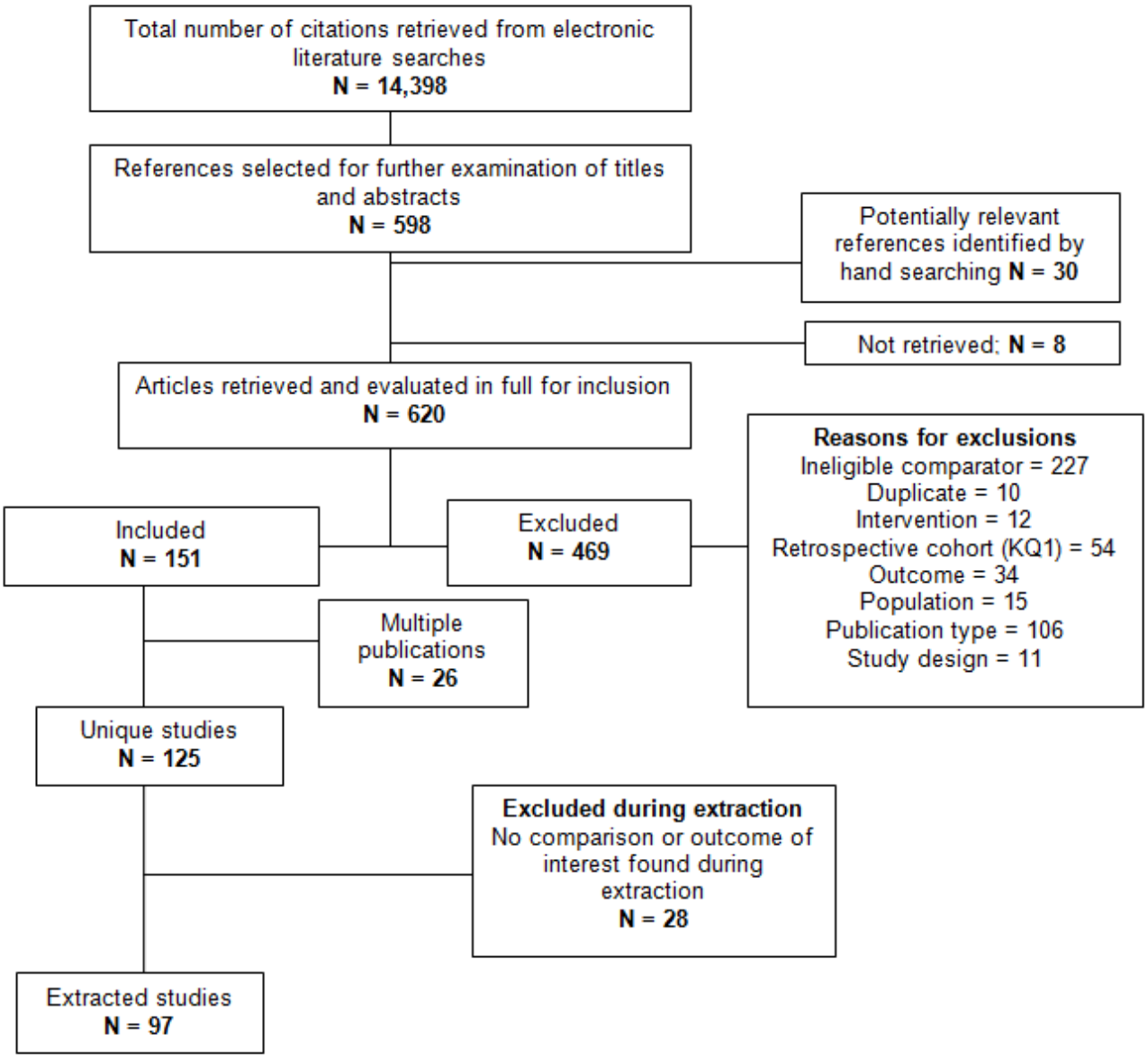
Several appendixes provide supporting information to the findings presented in this section. Appendix C provides the quality assessment ratings by domain for each study. Appendix D contains detailed evidence tables describing the study, characteristics of the population, screening criteria or diagnostic tests used, details of treatment (where relevant), and outcomes. A list of citations for the excluded and unobtained studies is available in Appendix E. Appendixes are available at the Agency for Healthcare Research and Quality (AHRQ) Web site www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Results of Literature Searches

The search strategy identified 14,398 citations from electronic databases. Screening based on titles and abstracts identified 598 potentially relevant studies. We identified 30 additional studies by hand searching the reference lists from included studies. Using the detailed selection criteria, 151 studies met the inclusion criteria and 469 were excluded. Of the 151 studies, 26 were identified as companion publications and 125 were unique studies (Figure 3). Of the 125 unique studies, 28 were further excluded during data extraction due to a lack of comparison or outcome of interest, leaving the total number of included studies at 97.

The most frequent reasons for exclusion were: (1) ineligible comparator (studies comparing two or more treatments but lacking a control group; n = 227); (2) ineligible publication type (abstracts, conference proceedings, studies published prior to 1995; n = 106); (3) ineligible study design (studies other than randomized controlled trials [RCTs], nonrandomized controlled trials [NRCTs], prospective cohort studies, and retrospective cohort studies; n = 11); (4) study did not report prespecified outcomes of interest (lacking test properties for Key Question 1, specified outcomes for Key Questions 3,4, and 5 including harms of screening or treatment; n = 34); (5) duplicate publication (n = 10); (6) intervention not of interest (studies without evaluation of screening tests or criteria, or treatments for gestational diabetes mellitus [GDM]; n = 12); and (7) population was not of interest (if >20 percent of pregnant women enrolled in study had known pre-existing diabetes without subgroup analysis; n = 15). In addition, for Key Question 1 only prospective studies were eligible for inclusion; 54 retrospective cohort studies were excluded. A complete list of excluded studies and reasons for exclusion is provided in Appendix E.

Figure 3. Flow diagram of study retrieval and selection



Key Question 1	51
Key Question 2	2
Key Question 3	38
Key Questions 4 and 5	11
TOTAL	**102**

* Five studies addressed more than one Key Question, therefore the sum of studies addressing the Key Questions exceeds the total number of studies.

Description of Included Studies

A total of 97 studies met the eligibility criteria for this review, including 6 RCTs, 63 prospective cohort studies, and 28 retrospective cohort studies. The studies were published between 1995 and 2012 (median 2004). Studies were conducted in the United States (24 percent), Europe (23 percent), Asia (22 percent), the Middle East (20 percent), Australia (4 percent), Central and South America (3 percent), and Canada (4 percent). The source of funding for the included studies was academic (23 studies, 24 percent), foundation or organization (17 studies, 18 percent), government (14 studies, 14 percent), "other" (such as the WHO, or non-governmental organization; 8 studies, 10 percent), and industry (10 studies, 10 percent). Twenty-two studies presented more than one source of funding. Two studies reported no external source of funding (2 percent), and 46 studies (47 percent) did not describe a source of funding.

Forty-eight studies (50 percent) analyzed women tested for GDM between 24-28 weeks, with a OGCT taking place first and the OGTT following within 7 days.[50,62-108] Thirty-one studies (32 percent) did not specify when screening or diagnostic procedures took place.[54,109-137] Of the 31 studies, one scheduled testing between 24 and 28 weeks, with different undefined test points if clinically warranted.[138] Eighteen studies (18 percent) screened or tested within unique time ranges.[133,139-155] Of these, one study screened participants with a OGCT at 21-23 weeks followed by a diagnostic OGTT at 24-28 weeks;[140] another screened a group of participants after 37 weeks;[146] one study screened before 24 weeks; [143] and one study screened women at risk between 14-16 weeks with normal women screened at the usual 24-28 weeks.[148] Remaining studies generally provided broader screening times ranging from 21-32 weeks gestation.[139,142,144,145,150-152] Studies employing WHO criteria generally screened further into gestation as only an OGTT was performed: one study screened at 28-32 weeks,[149] one study between 26-30 weeks,[155] another between 25-30 weeks,[154] and another study screened women at high risk at 18-20 weeks and others at 28-30 weeks.[147] One study using WHO criteria did not specify the time of testing.[133]

The number of women enrolled in each study ranged from 32[143] to 23,316[3] (median 750). The mean age of study participants was 30 years. The mean age was consistent among most studies, although women of slightly younger mean age (23-28 years) were enrolled studies originating from countries outside North America (India, Turkey, Hong Kong, United Arab Emirates).[113,114,144,156]

When duration of followup was reported, it was often described as "until birth" or "to delivery."[62,73,84,95,114,120,146,152] One study reported followup extending from the first prenatal visit (<13 weeks) until a OGCT (26-29 weeks),[139] one study within the first trimester until 24-28 weeks gestation,[101] and another began at first antenatal booking which ranged from first trimester through to the third in women who were present for antenatal care in late gestation.[157] One study followed women for 3 months postpartum;[83] and two studies provided longer-term followup extending to 5-7 years[132] and 7-11 years, respectively.[96] Remaining studies did not provide specific details on duration of followup.

Methodological Quality of Included Studies

The methodological quality of each study was assessed by two independent reviewers. Our approach to assessing study quality is described in the methods section. The consensus ratings for each study and domains are presented in Appendix C, Tables C1, C2, and C3. Studies were assessed using different tools depending on the Key Question and study design: for Key

Question 1, QUADAS-2 was used; for Key Questions 2 to 5, the Cochrane Risk of Bias tool was used for RCTs and the Newcastle Ottawa Scale was used for cohort studies. The methodological quality of studies is described in detail within the results section for each Key Question.

Key Question 1. What are the sensitivities, specificities, reliabilities, and yields of current screening tests for GDM?

GDM is diagnosed by having one or several glucose values at or above set glucose thresholds following an OGTT administered in the fasting state during pregnancy. Variations in glucose dose, time intervals of glucose measurements, and diagnostic glucose threshold values exist (Table 1). The most commonly used screening practice is a 50 g OGCT without regard to timing of last meal; plasma glucose is measured 1-hour after the glucose challenge. This was first proposed by O'Sullivan and Mahan[158] and has been modified over the years. There are two different glucose threshold values commonly used for this screen in North America: ≥140 mg/dL (≥7.8 mmol/L) and ≥130 mg/dL (≥7.2 mmol/L). Clinical and historical risk factors and fasting plasma glucose (FPG) are two other screening practices included in this current review.

Two related issues make it difficult to organize and analyze the studies that address Key Question 1. First, there are several screening options (e.g., risk factor-based, universal), and several techniques (e.g., glucola-based, fasting, postprandial). In addition, there is no 'gold standard' for diagnosing GDM. There are five different, but commonly used, glucose-based diagnostic measures that overlap in the criteria they use.

We grouped studies according to the comparator OGTT diagnosis practices that were used, specifically glucose load, time intervals, and threshold values. These groupings include: 3-hour, 100 g OGTT using Carpenter and Coustan (CC) criteria; 3-hour, 100 g OGTT using National Diabetes Data Group (NDDG) criteria; 2-hour, 75 g OGTT using American Diabetes Association (ADA) (2000-2010) criteria, and, 2-hour, 75 g OGTT using WHO criteria (Table 1). We present results of screening tests based on these groupings that included women who underwent the 50 g OGCT screen (further subdivided by screening threshold ≥140 mg/dL and ≥130 mg/dL), fasting plasma glucose (FPG), clinical and historical risk factors, and other screening criteria. This is followed by a section on studies that compared early and late screening practices. The final section summarizes the evidence comparing different glucose loads for the OGTT diagnostic tests. Forest plots present 2x2 data, sensitivity and specificity; summary tables present prevalence, positive and negative predictive values (PPV, NPV), and accuracy for individual studies.

Description of Included Studies

There were 51 studies (reported in 52 papers) that met the inclusion criteria for Key Question 1.[62-77,91,99-101,104,105,107-115,117-121,123-127,129,138-140,142-144,151,153,157] Two papers from the Tri-Hospital group[142] are included as they report on results for different screening practices.[159,160] Studies were conducted in a wide range of regions: 11 in North America,[64,69,72,104,105,121,123,126,127,142,143] 10 in Europe,[62,65,66,68,108,115,119,125,151,153] 12 in Asia,[70,73,101,107,111,114,118,128,129,139,140,157] 15 in the Middle East,[67,71,74-77,99,100,109,110,112,113,117,138,144] 2 in South America,[63,120] and 1 in Australia.[124] All studies were prospective cohort studies. A summary table of the study and patient characteristics of the individual studies can be found in Appendix D.

The prevalence of GDM varied across studies. The variability is due to differences in study setting (i.e., country), screening practices (e.g., universal vs. selective), and/or population characteristics (e.g., race/ethnicity, age, body mass index [BMI], parity). The range of GDM

prevalence for each diagnostic criteria is as follows: CC/ADA (2000-2010) (100 g) 3.6 to 38.0 percent; National Diabetes Data Group (NDDG) 1.4 to 50.0 percent, ADA (2000-2010) (75 g) from 2.0 to 19.0 percent, and WHO from 1.7 to 24.5 percent. Prevalence results for individual studies are presented in the following sections.

Methodological Quality of Included Studies

We used the QUADAS-2 tool to assess the quality of the studies included in this review. The tool comprises four key domains that discuss patient selection, index test, reference standard, and flow of patients through the study and the timing of the index tests and reference standard (flow and timing). The first part of QUADAS-2 concerns bias; the second part considers applicability or concerns that the study does not match the review question. Figure 4 summarizes the assessments for risk of bias and Figure 5 summarizes assessments of applicability. Detailed assessments for each study are presented in Appendix C1.

The domain of patient selection was rated as low risk that the selection of patients introduced bias for 53 percent of the studies. These studies were prospective cohort studies, most enrolled a consecutive sample of patients, and most avoided inappropriate exclusions. However, 25 percent of studies were rated as unclear due to inadequate description. Overall, 55 percent of studies were assessed as having high concerns about applicability for this domain. This was primarily because these studies were conducted in developing countries and used the WHO criteria to diagnose GDM. The results of these studies may not be directly relevant to the population in the United States.

The domain of the index test was generally rated as low risk that the conduct or interpretation of the index test introduced bias (53 percent). For most studies, the screening test (i.e., the index test) was conducted before the reference standard, and the threshold for the screening test was pre-specified. Concern about applicability was assessed as low (82 percent).

The domain of the reference standard (i.e., the criteria used to confirm a diagnosis of GDM) was generally rated as unclear risk that the conduct or interpretation of the reference standard introduced bias (63 percent). For most studies the result of the screening test was used to determine whether patients underwent further testing for GDM. Concern about applicability was assessed as low (86 percent).

The domain of flow and timing was assessed as low risk of bias for 39 percent of the studies. For most studies, the interval between the index test and reference standard was appropriate according to the criteria used in the study. Most patients received the reference standard, and received the same reference standard. However, in 35 percent of studies not all patients received a confirmatory reference standard if the screening test was below a certain threshold. These were assessed as unclear risk of bias.

Figure 4. QUADAS-2 assessment of risk of bias by domain

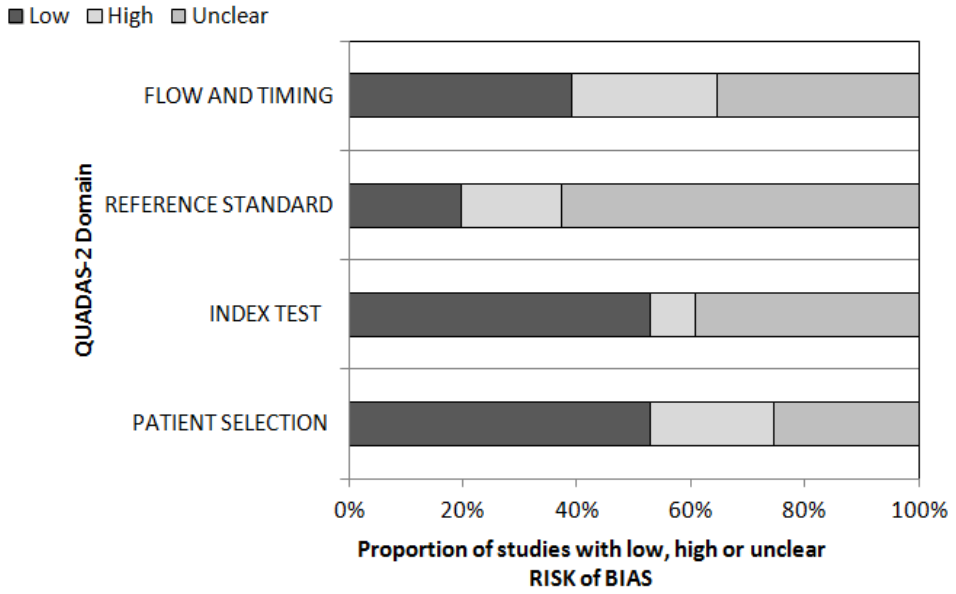
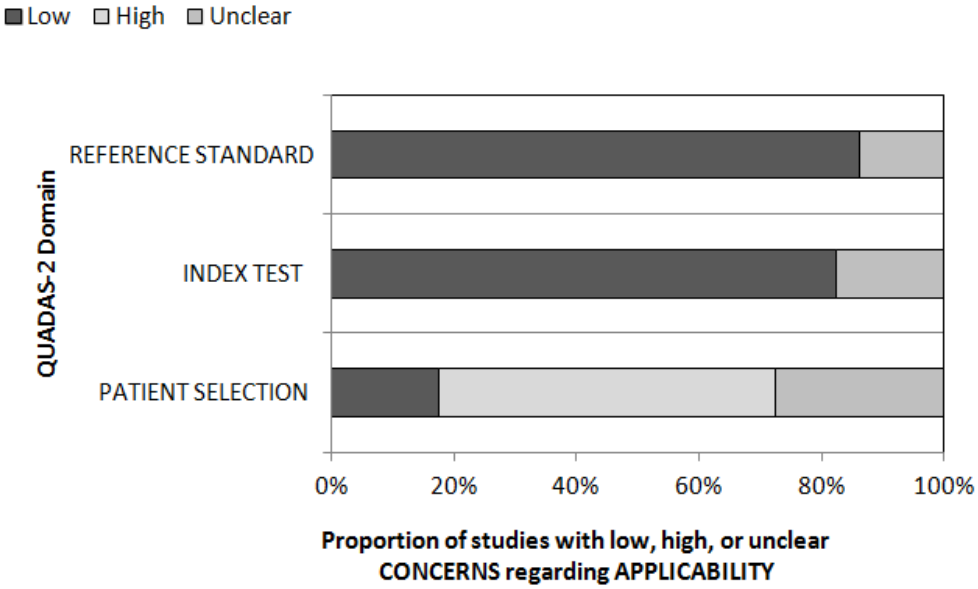


Figure 5. QUADAS-2 assessment of applicability by domain



Key Points

- Comparisons between screening tests and diagnostic thresholds were difficult because of the variety of different populations and different tests that were studied.
- Prevalence of GDM varied across studies and the diagnostic criteria used. The range of prevalence was: CC 3.6 to 38.0 percent; NDDG 1.4 to 50.0 percent; ADA (75 g) 2.0 to 19.0 percent; and WHO 1.7 to 24.5 percent.
- The 50 g OGCT with the 130 mg/dL cutpoint has higher sensitivity when compared with the 140 mg/dL cutpoint, however, specificity is lower (6 studies). Both thresholds have high NPV but variable PPV across a range of GDM prevalence.
- The use of a high cutoff for a diagnosis of GDM on an OGCT is supported by one study that assessed a 50 g OGCT (≥200 mg/dL) with GDM confirmed using the CC criteria. Sensitivity, specificity, PPV and NPV were all 100 percent.
- Fasting plasma glucose at a threshold of ≥85 mg/dL has similar sensitivity to 50 g OGCT; specificity is lower (4 studies).
- There were sparse data to assess screening and diagnostic tests for GDM less than 24 weeks' gestation.
- Four studies compared a 75 g load with a 100 g load (reference standard) to diagnose GDM. The prevalence of GDM ranged from 1.4 to 50 percent. Median sensitivity and PPV were low; median specificity and NPV were high.
- One study compared the IADPSG criteria with a two-step strategy. Sensitivity was 82 percent and specificity was 94 percent. Prevalence of GDM was 13.0 percent with IADPSG criteria compared with 9.6 percent with the two-step strategy. PPV and NPV were 61 percent and 98, respectively.

Detailed Synthesis

50 g OGCT Screening and GDM Diagnosis with 100 g OGTT

This section includes studies in which women underwent a 2-step practice that included screening with a 50 g OGCT at 24 to 28 weeks followed by a 100 g OGTT to confirm a diagnosis of GDM. The 50 g OGCT screening test is grouped by the two following diagnostic confirmation criteria: CC and ADA (2000-2010) criteria and the NDDG criteria.

Carpenter and Coustan and ADA (2000-2010) Criteria

Description of Included Studies

Fourteen studies confirmed a diagnosis of GDM with a 100 g, 3-hour OGTT using CC/ADA 2000-2010 criteria (Appendix D).[63,64,68,72,75-77,99,104,108,121,140,159,161] Ten studies used a universal screening practice,[63,64,68,72,76,77,108,121,159,161] three studies used a selective, risk-based screening practice for an OGCT,[75,99,140] and one study only included women with an abnormal OGCT.[104] Six studies performed the OGTT on all women regardless of OGCT value,[63,68,72,108,140,159] while eight performed an OGTT in patients with a positive OGCT.[64,75-77,99,104,121,161]

Studies were conducted in the United States,[64,104,121] Canada,[15] Iran,[71,75-77] Brazil,[63] France,[108] Mexico,[72] Switzerland,[68] Thailand,[140] and United Arab Emirates.[99] The number of patients analyzed ranged from 138 to 11,545. Maternal age was reported in 12 studies and the mean

ranged from 23.7 to 32.5 years. Mean BMI was reported in 10 studies and ranged from 23.3 to 29.6 kg/m^2. One study included women tested at ≥20 weeks' gestation.[121]

Results

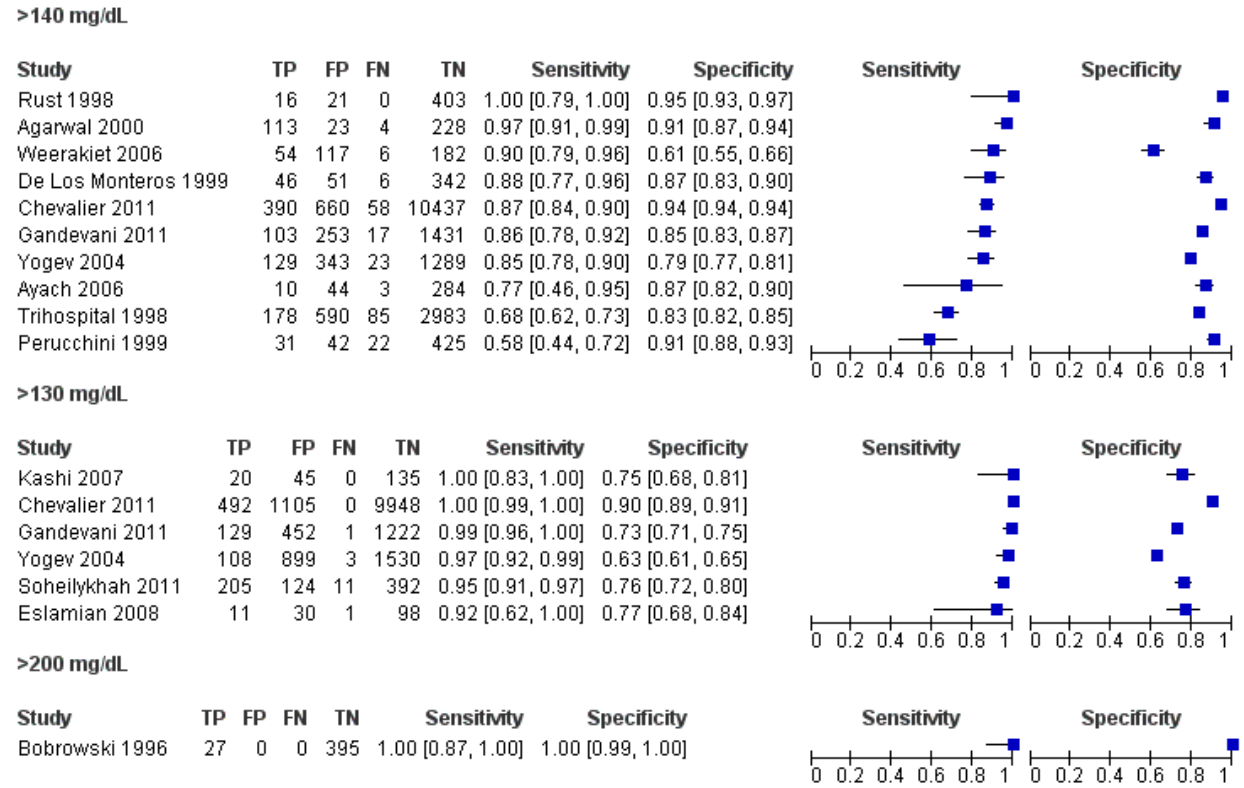
Nine studies provided data to estimate the test characteristics of a 50 g OGCT screening tested at the 1-hour interval and cutoff value of ≥140 mg/dL.[63,64,68,72,76,99,108,140,159] The accuracy of the OGCT (i.e., the proportion of true positive and true negative results) was generally high (median = 86.5 percent) and ranged from 66 to 94 percent (Table 3). Figure 6 presents the sensitivities and specificities for the individual studies. The joint estimates of sensitivity and specificity were 85 percent (95% CI, 76 to 90) and 86 percent (95% CI, 80 to 90). Hierarchical summary receiver operator characteristic (HSROC) curves comparing the sensitivity and specificity for all studies are presented in Appendix F. The prevalence of GDM ranged from 3.8 to 31.9 (Table 3). The PPV ranged from 18.5 to 83.1 percent; the NPV ranged from 95.1 to 99.0 percent (Table 3). The study by Rust et al. [121] included women ≥20 weeks and reported a sensitivity of 56 percent (95% CI, 30 to 80) and specificity of 94 percent (95% CI, 91 to 96). The prevalence of GDM was 3.6 percent.

Six studies used an OGCT cutoff value of ≥130 mg/dL.[64,71,75-77,108] The accuracy of the OGCT ranged from 64.5 to 90.4 (median = 78.5 percent) (Table 3). Figure 6 presents the sensitivities and specificities for the individual studies. The joint estimates of sensitivity and specificity were 99 percent (95% CI, 95 to 100) and 77 percent (95% CI, 68 to 83), respectively. The prevalence of GDM ranged from 4.3 to 29.5 (Table 3). The PPV ranged from 10.7 to 62.3 percent; the NPV ranged from 97.3 to 100 percent (Table 3).

One study used an OGCT cutoff value of >200 mg/dL.[104] The prevalence was 29.4 percent. The sensitivity was 100 (95% CI, 0.87 to 100) and specificity was 100 percent (95% CI, 0.99 to 100).

The studies by Agarwal,[99] Weerakiet,[140] Bobrowski, [104] and Kashi[75] are at high risk for selection bias due to the use of selective screening practice. Not all women received a confirmatory OGTT in the studies by Eslamian,[71] Gandevani,[76] Soheilykhah,[77] and Yogev[64] are at high risk for partial verification bias.

Figure 6. Forest plot of sensitivity and specificity: 50 g OGCT by CC or ADA (2000–2010) criteria



ADA = American Diabetes Association; CC = Carpenter-Coustan; FN = false negative; FP = false positive; OGCT = oral glucose challenge test; TN = true negative; TP = true positive

Table 3. Prevalence and diagnostic test characteristics for 50 g OGCT by CC or ADA (2000–2010) diagnostic criteria

Diagnostic Test	Author, Year	Country	N*	Screening Practice**	Prevalence (%)	PPV (95% CI)	NPV (95% CI)	Accuracy (%)
≥140 mg/dL OGCT	Rust, 1998[121]	U.S.	448	Universal	3.6	24 (13-40)	98 (97-99)	92
	Ayach, 2006[63]	Brazil	341†	Universal	3.8	18 (10-31)	99(97-100)	86
	Chevalier, 2011[108]	France	11,545†	Universal	3.9	37 (34-40)	99 (99-100)	94
	Trihospital, 1998[159]	Canada	3,836†	Universal	6.9	23 (20-26)	97 (96-98)	82
	Yogev 2004[64]	U.S.	1,783	Universal	8.5	27 (24-32)	98 (97-99)	80
	Perucchini, 1999[68]	Switzerland	520†	Universal	10.2	43 (32-54)	95 (93-97)	88
	De los Monteros, 1999[72]	Mexico	445†	Universal	11.7	47 (38-57)	98 (96-99)	87
	Weerakiet, 2006[140]	Thailand	359†	Selective	16.7	32 (25-39)	97 (93-99)	66
	Gandevani, 2011[76]	Iran	585	Universal	22.2	62 (55-69)	96 (93-97)	85
	Agarwal 2000[99]	UAE	368	Selective	31.9	83 (80-89)	98 (96-99)	93
≥130 mg/dL OGCT	Chevalier, 2011[108]	France	11,545†	Universal	4.3	31 (29-33)	100 (100-100)	90
	Yogev 2004[64]	U.S.	2,541	Universal	4.4	11 (9-13)	100 (99-100)	65
	Eslamian, 2008[71]	Iran	138	Universal	8.6	27 (16-42)	99 (95-100)	78
	Kashi, 2007[75]	Iran	200	Selective	10.0	31 (21-43)	100 (98-100)	78
	Gandevani, 2011[76]	Iran	585	Universal	22.2	51 (45-57)	100 (99-100)	79
	Soheilykhah, 2011[77]	Iran	1,502	Universal	29.5	62 (57-67)	97 (95-98)	82
≥200 mg/dL OGCT	Bobrowski, 1996[104]	U.S.	422†	Abnormal screen**	6.4	100 (91-100)	100 (100-100)	100 (99-100)

CI = confidence interval; NPV = negative predictive value; OGCT = oral glucose challenge test; OGTT = oral glucose tolerance test, PPV = positive predictive value; UAE = United Arab Emirates

*Number of women in the analysis.

**As reported in the methods of each study; all studies are 2-step screening and diagnosis.

†All women received both an OGCT and OGTT.

NDDG Criteria

Description of Included Studies

Ten studies used the NDDG criteria to confirm a diagnosis of GDM (Appendix D).[66,67,69,72-74,104,123,144,159] Eight studies used a universal screening practice;[66,67,69,72-74,144,159] two included only women with an abnormal OGCT.[104 123] Six studies performed the OGTT on all women regardless of OGCT value,[63,68,72,108,140,159] while the remaining studies performed an OGTT only in patients with a positive OGCT.

Four studies were conducted in North America,[69,104,123,159] two in Europe,[74,144] and one each in Mexico,[72] Saudi Arabia,[67] and Thailand,[73] and Turkey.[66] The number of patients enrolled ranged from 80 to 4,274. Mean maternal age, reported in seven studies, ranged from 25.7 to 32.1 years. Only two studies reported BMI. All studies screened women after 24 weeks' gestation.

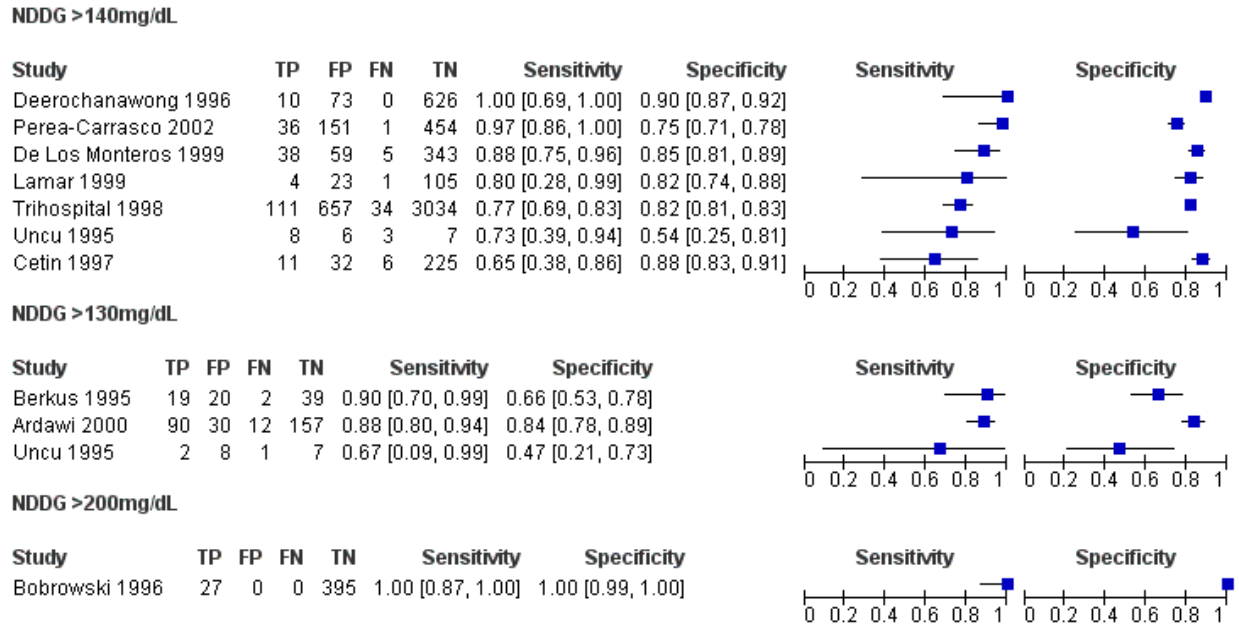
Results

Seven studies provided data to estimate the test characteristics of a 50 g OGCT tested at the 1-hour interval and cutoff value of ≥140 mg/dL.[66,69,72-74,144,159] The accuracy of the OGCT was generally high (median = 82 percent) (Table 4). Figure 7 presents the sensitivities and specificities for the individual studies. HSROC curves comparing the sensitivity and specificity for all studies are presented in Appendix F. The joint estimates of sensitivity and specificity were 85 percent (95% CI, 73 to 92) and 83 percent (95% CI, 78 to 87), respectively. The prevalence of GDM ranged from 1.4 to 45.8 (median = 6.2) (Table 4). The PPV ranged from 12.0 to 57.1; the NPV ranged from 70 to 100 (Table 4).

Three studies[67,74,113] used a cutoff ≥130 mg/dL. The accuracy of the test ranged from 50.0 to 85.5 percent (Table 4). Figure 7 presents the sensitivities and specificities for the individual studies. As there were only three studies, we did not pool the results. The prevalence of GDM ranged from 16.7 to 35.3 (Table 4). The PPV ranged from 20.0 to 75.0; the NPV ranged from 87.5 to 92.9 (Table 4). One study used an OGCT cutoff value of >200 mg/dL. The sensitivity, specificity, PPV and NPV were all 100 percent.

The studies by Ardawi,[67] Bobrowski,[104] Berkus[123] Cetin,[144] Deerochanawong, [73] Lamar,[69] and Uncu,[74] are at high or unclear risk for selection bias due to selective or unclear screening practices. Studies by Ardawi,[67] De los Monteros,[72] and Lamar,[69] are at high or unclear risk for partial verification bias as not all women received a confirmatory OGTT.

Figure 7. Forest plot of sensitivity and specificity: 50 g OGCT by NDDG criteria



NDDG = National Diabetes Data Group; OGCT = oral glucose challenge test

Table 4. Prevalence and diagnostic test characteristics for 50 g OGCT by NDDG diagnostic criteria

Diagnostic Test	Author, Year	Country	N*	Screening Practice**	Prevalence (%)	PPV (95% CI)	NPV (95% CI)	Accuracy (%)
≥140 mg/dL OGCT	Deerochanawong, 1996[73]	Thailand	709	Universal	1.4	12 (7-21)	100 (99-100)	90
	Trihospital, 1998[159]	Canada	3,836†	Universal	3.8	15(12-17)	99 (98-99)	82
	Lamar,1999[69]	U.S.	136	NR	3.8	15 (6-33)	99 (95-100)	82
	Perea-Carrasco, 2002[66]	Spain	578	Universal	5.8	19 (14-26)	100 (99-100)	76
	Cetin, 1997[144]	Turkey	274	Universal	6.2	26 (15-40)	97(95-99)	86
	De Los Monteros, 1999[72]	Mexico	445†	Universal	9.7	39 (30-49)	99 (97-99)	86
	Uncu, 1995[74]	Turkey	24†	Universal	45.8	57 (33-79)	70 (42-81)	63
≥130 mg/dL OGCT	Berkus, 1995[123]	U.S.	80†	NR	26.3	49 (34-64)	95 (85-98)	73
	Uncu, 1995[74]	Turkey	18†	Universal	16.7	20 (6-51)	86 (56-96)	50
	Ardawi, 2000[67]	Saudi Arabia	818	Universal	35.3	75 (67-82)	93 (88-96)	86
≥200 mg/dL OGCT	Bobrowski, 1996[104]	U.S.	422†	Abnormal screen	6.4	100 (91-100)	100 (100-100)	100

CI = confidence interval; NDDG = National Diabetes Data Group; NPV = negative predictive value; NR = not reported; OGCT = oral glucose challenge test; OGTT = oral glucose tolerance test; PPV = positive predictive value
*Number of women in the analysis.
**As reported in the methods of each study; all studies are 2-step screening and diagnosis.
†All women received both an OGCT and OGTT.

50 g OGCT Screening and GDM Diagnosis with 75 g OGTT

This section includes studies in which women underwent a 2-step screening and diagnostic practice that included a 50 g OGCT followed by a 75 g OGTT to confirm a diagnosis of GDM.

ADA (2000-2010) Criteria

Description of Included Studies

Three studies[101,125,139] used the ADA 75 g, 2-hour criteria after a 50 g, 1-hour OGCT (Appendix D). All but the study by Maegawa et al.[101] used a threshold of ≥140 mg/dL for the OGCT. The studies were conducted in Japan,[101,139] and Germany.[125] One Canadian study[105] confirmed diagnosis using the Canadian Diabetes Association 75 g, 2-hour criteria.

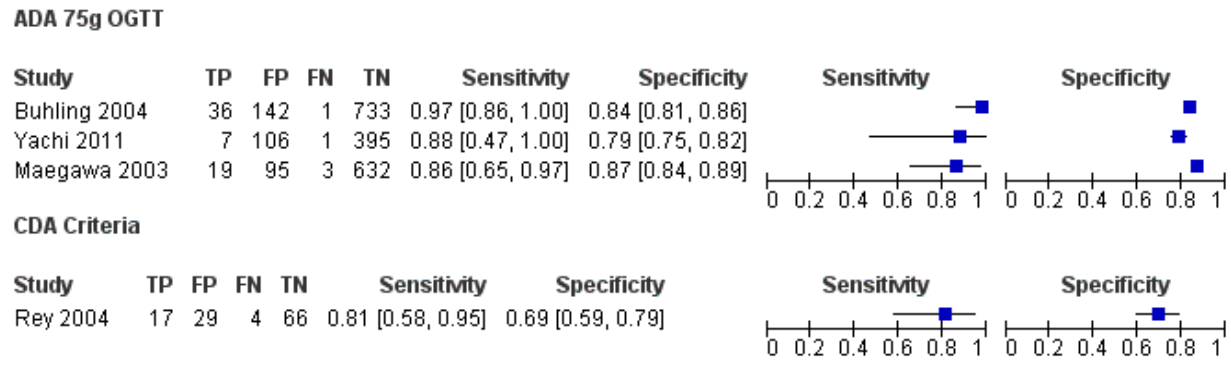
The number of patients analyzed ranged from 509 to 912. All studies reported maternal age, which ranged from 28.5 to 33.4 years. BMI ranged from 20.0 to 24.8 kg/m^2. All studies performed the OGCT screening at 24-28 weeks; two studies also screened women in early pregnancy.[101,139]

Results

The accuracy of the ADA (2000-2010) 75 g ranged from 84 percent to 87 percent (Table 5). Figure 8 presents the sensitivities and specificities for the individual studies. The results were not pooled. The prevalence of GDM ranged from 1.6 to 18.1 (Table 5). The PPV ranged from 7 to 20; the NPV ranged from 99 to 100 (Table 5). The accuracy of the CDA 75 g was 72 percent; PPV was 37 percent and NPV was 94 percent, respectively.

The studies by Rey[105] and Yachi[139] are at high or unclear risk of selection bias due to their screening practices. The study by Buhling,[125] is at high risk for partial verification bias as not all women received a confirmatory OGTT.

Figure 8. Forest plot of sensitivity and specificity: 50 g OGCT (different thresholds) by ADA (2000–2010) 75 g criteria



ADA = American Diabetes Association; CDA = Canadian Diabetes Association; OGCT = oral glucose challenge test; OGTT = oral glucose tolerance test

Table 5. Prevalence and diagnostic test characteristics for 50 g OGCT (different thresholds) by ADA (2000–2010) 75 g criteria

Organization	Author, Year	Country	N*	Screening Practice**	Prevalence (%)	PPV (95% CI)	NPV (95% CI)	Accuracy (%)
ADA (2000-2010)	Yachi, 2011[139]	Japan	509	Universal	1.6	7 (4-13)	100 (99-100)	79
	Maegawa, 2003[101]	Japan	749	Universal	2.9	17 (11-25)	99 (98-100)	87
	Buhling, 2004[125]	Germany	912	Universal	4.1	20 (15-27)	100 (99-100)	84
CDA	Rey, 2004[105]	Canada	188†	Selective	18.1	37 (25-51)	94 (87-97)	72

ADA = American Diabetes Association; CDA = Canadian Diabetes Association; CI = confidence interval; NPV = negative predictive value; OGCT = oral glucose challenge test, OGTT = oral glucose tolerance test; PPV = positive predictive value
*Number of women in the analysis.
**As reported in the methods of each study; all studies are 2-step screening and diagnosis.
†All women received both an OGCT and OGTT.

World Health Organization Criteria

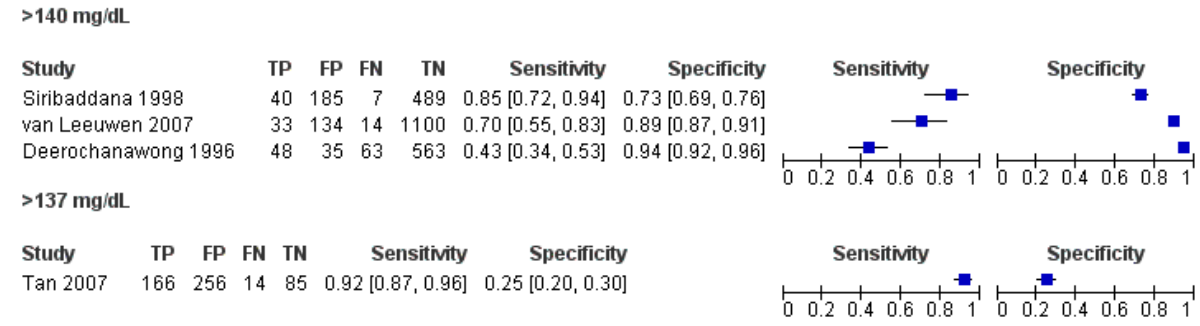
Description of Included Studies

Four studies used the WHO criteria to confirm a diagnosis of GDM (Appendix D).[62,70,73,157] The studies were conducted in Netherlands,[62] Sri Lanka,[70] Malaysia,[157] and Thailand.[73] The number of patients enrolled ranged from 188 to 1,301. Mean maternal age ranged from 25.7 to 30.8 years. Mean BMI, as reported in two studies, was 22.4 and 24.2. All studies performed the OGCT screening at 24-28 weeks with OGTT performed the following 1 to 2 weeks.

Results

The accuracy of the test ranged from 73 percent to 88 percent (Table 6). Figure 9 presents the sensitivities and specificities for the individual studies. The results were not pooled. The prevalence of GDM ranged from 3.7 to 15.7 (Table 6). The PPV ranged from 5 to 20; the NPV ranged from 94 to 99 (Table 6). The prevalence of GDM ranged from 3.7 to 50.0 (Table 6). The PPV ranged from 17.8 to 76.2; the NPV ranged from 78.9 to 98.7

Figure 9. Forest plot of sensitivity and specificity: 50 g OGCT by WHO criteria



OGCT = oral glucose challenge test; WHO = World Health Organization

Table 6. Prevalence and diagnostic test characteristics for 50 g OGCT by WHO diagnostic criteria

Diagnostic Test	Author, Year	Country	N*	Screening Practice**	Prevalence (%)	PPV (95% CI)	NPV (95% CI)	Accuracy (%)
≥140 mg/dL OGCT	van Leeuwen, 2007[62]	Netherlands	1,301	Universal	3.7	20 (14-26)	99 (98-99)	88
	Siribaddana, 1998[70]	Sri Lanka	721	Universal	6.5	18 (13-23)	99 (97-99)	73
	Deerochanawong, 1996[73]	Thailand	709	Universal	15.7	58 (47-68)	90 (87-92)	86
≥130 mg/dL OGCT	Tan, 2007[157]	Malaysia	521	Universal	34.6	39 (35-41)	86 (78-91)	48

CI = confidence interval; NPV = negative predictive value; OGCT = oral glucose challenge test; PPV = positive predictive value; WHO = World Health Organization

*Number of women in the analysis.

**As reported in the methods of each study; all studies are 2-step screening and diagnosis. Fasting Plasma Glucose Screening and GDM Diagnosis

This section includes studies that examined FPG as a screening test. A diagnosis of GDM was confirmed using CC or ADA (2000-2010), WHO, NDDG, and CDA 75 g OGTT criteria.

Fasting Plasma Glucose and CC/ADA (2000-2010) Criteria

Description of Included Studies

Seven studies provided data on FPG at various thresholds as an alternative screening test to glucola-based screening with a diagnosis of GDM using CC and ADA (2000-2010) criteria (Appendix D).[65,75,99,108,112,126,127] Three studies used a universal screening practice[112 108,127] and the remaining studies used a selective, risk-based screening practice.[65,75,99,126] All but one study[75] performed the OGTT on all women regardless of OGCT value.

Studies took place in the United States,[126,127] France,[65,108] Iran,[75] and the United Arab Emirates.[99,112] The number of patients enrolled ranged from 123 to 11,545. Mean maternal age was reported in four studies and ranged from 27.8 to 32.8 years. Mean BMI was reported in three studies and ranged from 22.5 to 29.6. Most studies tested women after 24 weeks' gestation; one study tested women at 23 weeks.[126]

Results

The studies provided data to estimate the test characteristics of FPG at four common thresholds: ≥85 mg/dL (4.7 mmol/L), ≥90 mg/dL (5.0 mmol/L), ≥92 mg/dL (5.1 mmol/L), and ≥95 mg/dL (5.3 mmol/L). Figure 10 presents the sensitivities and specificities for the individual studies. The joint estimates of sensitivity and specificity, respectively for the different FPG threshold values are:

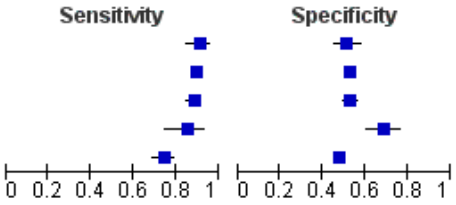
- ≥85 mg/dL: 87 percent (95% CI, 81 to 91) and 52 percent (95% CI, 50 to 55)
- ≥90 mg/dL: 77 percent (95% CI, 66 to 85) and 76 percent (95% CI, 75 to77)
- ≥92 mg/dL: 76 percent (95% CI, 55 to 91) and 92 percent (95% CI, 86 to 96) (median)
- ≥95 mg/dL: 54 percent (95% CI, 32 to 74) and 93 percent (95% CI, 90 to 96)

The prevalence of GDM ranged from 1.4 to 33.3 (median = 6.2) (Table 7). The PPV ranged from 12.0 to 45.8; the NPV ranged from 83.3 to 100 (Table 7).

Figure 10. Forest plot of sensitivity and specificity: fasting plasma glucose by CC/ADA (2000–2010) criteria

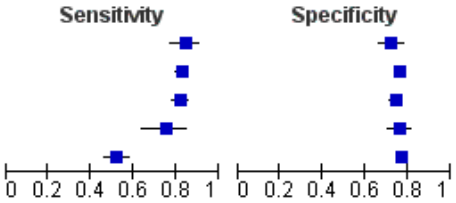
≥85 mg/dL

Study	TP	FP	FN	TN	Sensitivity	Specificity
Agarwal 2000	107	122	10	129	0.91 [0.85, 0.96]	0.51 [0.45, 0.58]
Agarwal 2006	539	1845	62	2081	0.90 [0.87, 0.92]	0.53 [0.51, 0.55]
Agarwal 2000	349	417	47	463	0.88 [0.85, 0.91]	0.53 [0.49, 0.56]
Kashi 2007	59	41	10	90	0.86 [0.75, 0.93]	0.69 [0.60, 0.77]
Sacks 2003	224	2018	78	1863	0.74 [0.69, 0.79]	0.48 [0.46, 0.50]



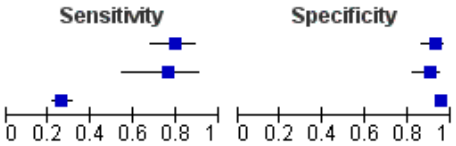
≥90 mg/dL

Study	TP	FP	FN	TN	Sensitivity	Specificity
Agarwal 2000	99	70	18	181	0.85 [0.77, 0.91]	0.72 [0.66, 0.78]
Agarwal 2006	497	939	104	2988	0.83 [0.79, 0.86]	0.76 [0.75, 0.77]
Agarwal 2000	323	224	71	658	0.82 [0.78, 0.86]	0.75 [0.72, 0.77]
Chastang 2003	52	68	17	217	0.75 [0.64, 0.85]	0.76 [0.71, 0.81]
Sacks 2003	157	964	145	3241	0.52 [0.46, 0.58]	0.77 [0.76, 0.78]



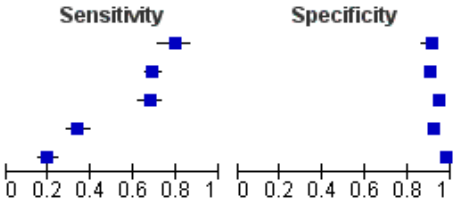
≥92 mg/dL

Study	TP	FP	FN	TN	Sensitivity	Specificity
Kashi 2007	55	10	14	120	0.80 [0.68, 0.88]	0.92 [0.86, 0.96]
Kauffman 2006	19	10	6	88	0.76 [0.55, 0.91]	0.90 [0.82, 0.95]
Chevalier 2011	87	51	243	1002	0.26 [0.22, 0.31]	0.95 [0.94, 0.96]



≥95 mg/dL

Study	TP	FP	FN	TN	Sensitivity	Specificity
Agarwal 2000	93	23	24	228	0.79 [0.71, 0.86]	0.91 [0.87, 0.94]
Agarwal 2006	414	399	187	3564	0.69 [0.65, 0.73]	0.90 [0.89, 0.91]
Agarwal 2000	219	53	105	827	0.68 [0.62, 0.73]	0.94 [0.92, 0.95]
Sacks 2003	102	341	200	3864	0.34 [0.28, 0.39]	0.92 [0.91, 0.93]
Chevalier 2011	65	24	265	1029	0.20 [0.16, 0.24]	0.98 [0.97, 0.99]



ADA = American Diabetes Association; CC = Carpenter-Coustan; OGCT = oral glucose challenge test

Table 7. Prevalence and diagnostic test characteristics for fasting plasma glucose by CC/ADA (2000–2010) diagnostic criteria

FPG by CC/ADA	Author, Year	Country	N*	Screening Practice**	Prevalence (%)	PPV (95% CI)	NPV (95% CI)	Accuracy (%)
FPG (≥85 mg/dL)	Agarwal, 2000[99]	UAE	1,276 (RF)	Selective	31.8	47 (40-53)	93 (87-96)	64
			398 (+OGCT)	Selective	31.0	46 (42-49)	91 (88-93)	64
	Agarwal, 2006[112]	UAE	4,609	Universal	13.3	23 (21-24)	97 (96-98)	58
	Kashi ,2007[75]	Iran	200	Selective	34.5	59 (49-68)	90 (83-94)	75
	Sacks, 2003[126]	U.S.	4,507	Universal	7.2	10 (9-11)	96 (95-97)	50
FPG (≥90 mg/dL)	Agarwal, 2000[99]	UAE	1,276 (RF)	Selective	31.8	59 (51-66)	91 (86-94)	76
			398 (+GCT)	Selective	30.9	59 (54-63)	91 (88-92)	77
	Agarwal, 2006[112]	UAE	4,609	Universal	13.3	35 (32-37)	97 (96-97)	77
	Chastang, 2003[65]	France	354	High risk	19.5	43 (34-52)	93 (89-95)	76
	Sacks, 2003[126]	U.S.	4,507	Universal	6.7	14 (12-16)	96 (95-96)	75
FPG (≥92 mg/dL)	Chevalier, 2011[108]	France	11,454	Universal	23.9	63 (55-71)	81 (78-83)	79
	Kashi, 2007[75]	Iran	200	Selective	34.7	85 (74-91)	90 (83-94)	88
	Kauffman, 2006[127]	U.S.	123	Universal	20.3	66 (47-80)	94 (87-97)	87
FPG (≥95 mg/dL)	Agarwal, 2000[99]	UAE	1,276 (>1 RF)	Selective	31.8	80 (72-86)	91 (86-93)	87
			398 (+OGCT	Selective	26.9	81 (75-85)	89 (87-90)	87
	Agarwal, 2006[112]	UAE	4,609	Universal	13.2	51 (48-54)	95 (94-96)	87
	Kashi , 2007[75]	Iran	200	Selective	23.9	73 (63-81)	80 (77-82)	79
	Sacks, 2003[126]	U.S.	4,507	Universal	6.7	23 (19-27)	95 (94-96)	88

ADA = American Diabetes Association; CC = Carpenter-Coustan; CI = confidence interval; FPG = fasting plasma glucose; NPV = negative predictive value; OGCT = oral glucose challenge test; PPV = positive predictive value; RF = risk factor screening; UAE = United Arab Emirates

*Number of women in the analysis.

**As reported in the methods of each study.

Fasting Plasma Glucose and Other Diagnostic Criteria

Description of Included Studies

Two studies used the WHO criteria to confirm a diagnosis of GDM,[111,120] one used the NDDG criteria,[127] and one each used the criteria from the national organizations from Canada[105] and Japan.[101] Different FPG thresholds were used: Maegawa et al.[101] and Wijeyaratne et al.[111] used ≥ 85 mg/dL, Kauffman et al.[127] used ≥ 92 mg/dL, and Reichelt et al.[120] used ≥ 89 mg/dL.

Results

Table 8 summarizes the prevalence and test characteristics of the studies.

Table 8. Prevalence and diagnostic test characteristics for fasting plasma glucose by NDDG-WHO and other diagnostic criteria

Criteria	Author, Year, Country	N*	Preva-lence (%)	Sn (%) (95% CI)	Sp (%) (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy (%)
WHO criteria	Reichelt,1998, Brazil[120]	4,977	0.3	88 (62-98)	78 (77-79)	1.3 (0.8-2.1)	100	78
	Wjieyaratne, 2006, Sri Lanka**[111]	853	16.9	92 (87-96)	71 (68-75)	40 (35-45)	98 (96-99)	75
NDDG criteria	Kauffman-2006, U.S.[127]	123	13.0	81 (54-96)	88 (80-93)	50 (32-68)	97 (92-99)	87
Other diagnostic criteria	Maegawa, 2003, Japan[101]	749 (1st Tri) (2nd Tri)	 1.9 2.9	 71 (68-79) 77 (72-80)	 83 (78-87) 91 (86-94)	 7 (4-13) 20 (13-30)	 99 (98-100) 99 (98-100)	 82 90
	Rey, 2004, Canada* [105]	122	17.2	90 (70-99)	46 (36-56)	22 (14-31)	94 (82- 98)	42

CI = confidence interval; NDDG = National Diabetes Data Group; NPV = negative predictive value; PPV = positive predictive value; Sn = sensitivity; Sp = specificity; Tri = trimester; WHO = World Health Organization
*Number of women in the analysis.
** Selective screening practice.

Risk Factor-Based Screening and GDM Diagnosis

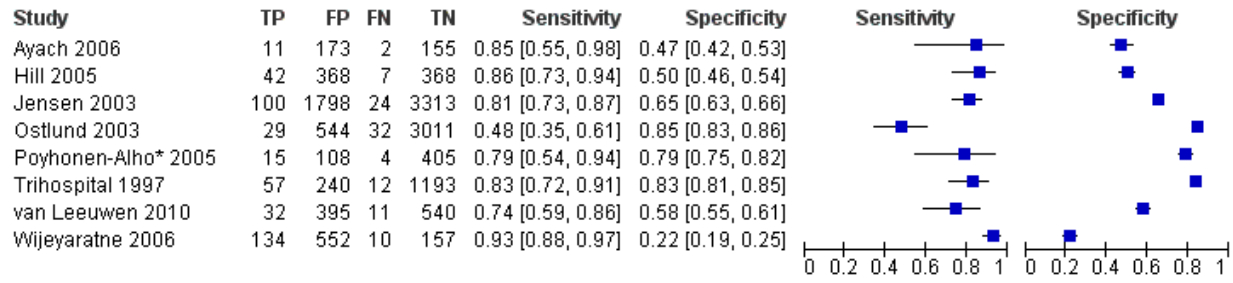
Description of Included Studies

Eight studies presented data on risk factor-based screening (Appendix D).[63,99,111,114,115,119,151,160] One study was conducted in North America,[160] four in Europe,[115,119,151,162] two in the Middle East,[114 111] and one in South America.[63] The number of patients enrolled ranged from 532 to 4,918.

Results

Figure 11 presents the sensitivities and specificities for the individual studies. The results were not pooled because different diagnostic criteria were used across the studies (Table 9). The prevalence of GDM ranged from 1.7 to 16.9 (Table 9). The PPV ranged from 5 to 20; the NPV ranged from 94 to 99 (Table 9).

Figure 11. Forest plot of sensitivity and specificity: risk factor screening by different diagnostic criteria (CC/ADA, NDDG, WHO)



ADA = American Diabetes Association; CC = Carpenter-Coustan; NDDG = National Diabetes Data Group; WHO = World Health Organization
*author-defined threshold values

Table 9. Prevalence and diagnostic test characteristics for risk factor screening by different diagnostic criteria

Criteria	Author, Year	Country	N*	Screening Practice**	# RF	Prevalence (%)	PPV (95% CI)	NPV (95% CI)	Accuracy (%)
CC/ADA (2000-2010)	Ayach, 2006[63]	Brazil	341	Universal	≥1	3.8	6 (3-10)	99 (96-100)	49
	Hill, 2005[114]	India	830	Universal	≥1	6.2	10 (8-14)	98 (96-99)	52
NDDG	Trihospital, 1997[160]	Canada	3,131	Universal	≥2	4.6	19 (15-24)	99 (98-99)	83
WHO	Ostlund, 2003[119]	Sweden	4,918	Universal	≥1	1.7	5 (4-7)	99 (99-100)	84
	Jensen, 2003[115]	Denmark	5,235	Universal	≥1	2.4	5 (4-6)	99 (98-100)	65
	Wijeyaratne, 2006[111]	Sri Lanka	853	Universal	≥1	16.9	20 (17-23)	94 (89-97)	34
Author-defined criteria	Poyhonen-Alho, 2005[151]	Finland	532	Universal	≥1	3.6	12 (8-19)	99 (98-100)	79

ADA = American Diabetes Association; CC = Carpenter-Coustan; CI = confidence interval; NDDG = National Diabetes Data Group; NPV = negative predictive value; PPV = positive predictive value; RF = risk factor; WHO = World Health Organization
*Number of women in the analysis.
**As reported in the methods of each study.

Other Screening Tests

Other studies examined point of care testing with a glucometer to measure capillary blood glucose,[110,111,116,117,128] or other markers such as fasting plasma insulin,[127,139] serum fructosamine,[74,109] glycated hemoglobin (HbA1c),[74,113] adiponectin levels,[140] and glycosuria.[125] The results are summarized in Table 10.

Table 10. Prevalence and characteristics of other screening tests by GDM diagnostic criteria

Screening Test	Author, Year Country	N*	Index Test Threshold	Reference Standard	Prevalence (%)	Sn (%) (95% CI)	Sp (%) (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy (%)
HbA1c	Uncu, 1995, Turkey[74]	42	7.2%	CC	33.3	64 (35-87)	64 (44-81)	47 (27-68)	78 (59-87)	64
	Agarwal, 2005, UAE[113]	442	7.5%	ADA (75 g)	19.0	82 (72-90)	21 (17-26)	20 (16-24)	83 (75-90)	33
	Agarwal 2001, UAE[100]	430	5.0%	CC	26.8	92 (86-96)	28 (23-33)	32 (27-37)	91 (83-95)	45
	Rajput, 2011, India[107]	607	5.5% 5.3%	ADA IADPSG	7.1 23.7	86 (72-95) 12 (7-18)	61 (57-65) 97 (95-98)	15 (11-19) 57 (39-73)	98 (96-99) 78 (74-82)	63 77
Serum fructosamine	Agarwal, 2011, UAE[109]	849	≥237 µmol/L	ADA (75 g)	13.3	86 (78-92)	23 (20-27)	15 (12-18)	92 (87-95)	32
	Uncu, 1995, Turkey[74]	42	≥2.85 mmol/L	CC	33.3	71 (42-92)	46 (28-66)	40 (23-59)	77 (55-86)	55
	Agarwal 2001, UAE[100]	430	≥210 µmol/L	CC	26.7	92 (86-96)	23 (18-28)	31 (26-36)	89 (81-94)	42
Fasting plasma insulin	Kauffman, 2006, U.S.[127]	123	≥93 µmol/L	NDDG	13.0	56.0 (35-76)	71 (61-80)	33 (21-48)	86 (78-92)	68
	Yachi, 2007, Japan[139]	509	≥3.66 mmol/L	JSOG (10 wk)	2.0	48 (43-53)	72 (63-80)	86 (80-90)	29 (24-36)	53
Author defined = (fructosamlne/ total protein) - (glucose/100)	Perea-Carrasco, 2002, Spain[66]	578	≥27.2	IWC, 3rd	7.0	98 (90-100)	89 (86-91)	44 (35-53)	100 (99-100)	90
Adiponectin	Weerakiet ,2006, Thailand[140]	359	≥10 µg/mL	ADA	16.7	92 (82-97)	31 (26 to 36)	18 (14-23)	96 (91-98)	40
Capillary blood glucose	Agarwal, 2008, UAE[110]	1,662	≥88 mg/dL	ADA (FPG)	11.2	84 (78-89)	75 (73-77)	30 (26-34)	98 (96-98)	76
	Balaji 2012, India[128]	819	≥140 mg/dL	WHO	10.5	80 (70-88)	98 (97-99)	86 (77-92)	98 (96-99)	97
	Wijeyaratne, 2006, Sri Lanka[111]	853	≥130 mg/dL	WHO	16.3	63 (54-70)	37 (34-41)	17 (14-20)	83 (79-87)	42
Glucose source	Eslamian, 2008, Iran[71]	138	50 g carb breakfast	ADA	8.6	83 (52-98)	86 (79-91)	36 (20-5)	98 (94-100)	86
	Lamar, 1999, U.S.[69]	136	50 g (28) jelly beans	NDDG	3.7	40 (5-85)	85 (78-91)	9 (3-28)	97 (93-99)	83
	Rust,1998, U.S.[121]	448	100 g carb meal	ADA (20 wk)	3.6	25 (7-52)	98 (96-99)	40 (17-69)	96 (93-98)	94

ADA = American Diabetes Association; carb = carbohydrate; CC = Carpenter-Coustan; CI = confidence interval; FPG = fasting plasma glucose; GDM = gestational diabetes mellitus; HbA1c = glycated hemoglobin; IADPSG = International Association of Diabetes in Pregnancy Study Groups; IWC = International Workshop Conference; JSOG = Japan Society of Obstetrics and Gynecology; NDDG = National Diabetes Data Group; NPV = negative predictive value; PPV = positive predictive value; Sn = sensitivity; Sp = specificity; UAE = United Arab Emirates; WHO = World Health Organization

*Number of women in the analysis

Comparison of Early and Late Screening Tests

One study (n = 749) conducted in Japan provided data on screening for GDM in the first and second trimesters.[101] The authors used three different screening tests: FPG, HbA1c, and a casual 50 g, 1-hour OGCT. GDM was confirmed with Japan Society of Obstetrics and Gynecology criteria (75 g, 2-hour) 2 to 4 weeks after screening. Prevalence of GDM using a universal screening practice was 1.9 percent in the first trimester and 2.9 percent in the second trimester. Table 11 presents a summary of the test characteristics by screening test and time point. These results should be interpreted cautiously as the women diagnosed with GDM in the first trimester had pre-pregnancy body weight and BMI that were significantly higher than for women who did not have GDM.

Table 11. Prevalence and characteristics of various screening tests for screening in the first and second trimesters (Maegawa study)

Screening Test	Trimester	Prevalence (%)	Sn (%)	Sp (%)	PPV (%)	NPV (%)
FPG (85 mg/dL)	First trimester	1.9	71.4	83.0	7.4	99.2
	Second trimester	2.9	77.0	90.7	20.0	99.3
50 g OGCT (threshold 130 mg/dL)	First trimester	1.9	92.9	77.0	7.1	99.8
	Second trimester	2.9	100.0	85.4	17.2	100
HbA1c (threshold 4.8%; 83.5% ULN)	First trimester	1.9	71.4	70.8	4.4	99.2
	Second trimester	2.9	36.4	72.9	3.9	97.4
HbA1c (threshold 5.8%)	First trimester	1.9	28.6	100	100	98.7
	Second trimester	2.9	13.6	99.9	75	97.4

FPG = fasting plasma glucose; HbA1c = glycated hemoglobin; NPV = negative predictive value; OGCT = oral glucose challenge test; PPV = positive predictive value; Sn = sensitivity; Sp = specificity; ULN = upper limit of normal

Comparison of Different Diagnostic Criteria

Seven studies provided data on the comparability of two diagnostic tests in the same group of women. The diagnostic criteria were: 75 g, 2-hour versus 100 g, 3-hour criteria; IADPSG versus the two-step Australasian Diabetes in Pregnancy Society (ADIPS) criteria; FPG versus ADA 100 g, 3-hour criteria; and IADPSG FPG ≥92 mg/dL versus WHO 75 g criteria.

Four studies compared 75 g, 2-hour criteria with 100 g, 3-hour criteria as the reference standard; however, different populations were assessed (Figure 12). The study by Brustman (n = 32) was conducted in the United States and compared the results of a 75 g, 3 hour OGTT with a 100 g, 3 hour OGTT.[143] Prevalence of GDM was 50 percent with NDDG criteria. The sensitivity was 29 percent (95% CI, 8 to 58) and the specificity was 89 percent (95% CI, 65 to 99); PPV and NPV were 100 (95% CI, 69 to 100) and 62 (95% CI, 43 to 72), respectively.

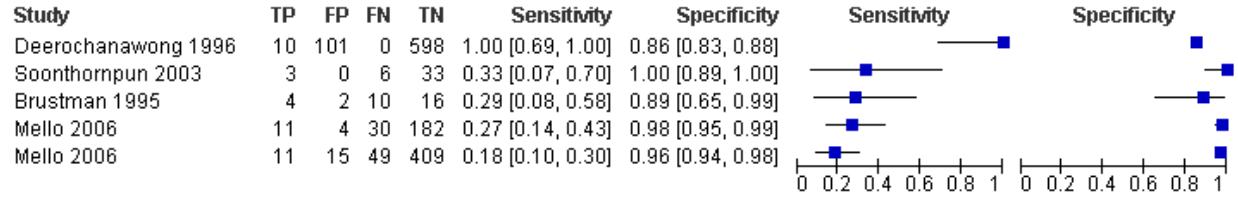
The study by Deerochanawong was conducted in Thailand (n = 709).[73] The prevalence of GDM was 1.4 percent with NDDG criteria and with WHO criteria it was 15.7 percent. Sensitivity was 100 percent (95% CI, 69 to 100) and specificity was 90 percent (95% CI, 92 to 96). PPV and NPV were 12 (95% CI, 7 to 21) and 100 (95% CI, 99 to100), respectively.

The study by Soonthornpun was also conducted in Thailand (n = 42).[118] The prevalence of GDM using the CC criteria was 21 percent. Sensitivity was 33 percent (95% CI, 7 to 70) and specificity was 100 percent (95% CI, 89 to 100). PPV and NPV were 100 (95% CI, 53 to 100) and 85 (95% CI, 71 to 92), respectively.

The fourth study by Mello was conducted in Italy and assessed diagnosis of GDM in women during early pregnancy (16 to 21 weeks) (n = 227) and late pregnancy (26 to 31 weeks) (n =

484).[153] For the early pregnancy group, the prevalence using CC criteria was 18 percent. Sensitivity was 27 percent (95% CI, 14 to 43) and specificity was 98 percent (95% CI, 95 to 99). PPV and NPV were 73 (95% CI, 48 to 89) and 86 (95% CI, 81 to 90), respectively. For the late pregnancy group the prevalence of GDM was 12 percent. Sensitivity was 18 percent (95% CI, 10 to 30) and specificity was 96 percent (95% CI, 94 to 98). PPV and NPV were 42 (95% CI, 25 to 61) and 89 (95% CI, 86 to 92), respectively.

Figure 12. Forest plot of sensitivity and specificity: 75 g OGTT by 100 g OGTT



OGTT = oral glucose tolerance test

An Australian study (n = 1,275) compared the diagnosis of GDM using IADPSG criteria with the ADIPS criteria as the reference standard.[124] GDM prevalence was 13.0 percent with IADPSG criteria compared with 9.6 percent with ADIPS. The sensitivity of IADPSG was 82 percent (95% CI, 74 to 88) and specificity was 94 percent (95% CI, 93 to 96); the PPV and NPV were 61 percent (95% CI, 53 to 68) and 98 (95% CI, 97 to 99), respectively.

Two studies assessed FPG as a diagnostic test but used different reference standards. A Brazilian study (n = 341) compared FPG with the ADA 100 g, 3-hour criteria.[63] The prevalence of GDM was 3.8 percent using ADA (2000-2010) 100 g criteria. The sensitivity was 84 percent (95% CI, 55 to 98) and specificity was 47 percent (95% CI, 42 to 53); PPV and NPV were 6 (95% CI, 3 to10) and 99 (95% CI, 56 to 100), respectively.

The second study, conducted in India (n = 1,463), compared IADPSG FPG criteria with the WHO 75 g criteria.[107] The prevalence of GDM was 13.4 percent with WHO criteria and 3.2 percent with FPG (≥95 mg/dL). The sensitivity of FPG as a diagnostic test was 29 percent (95% CI, 29 to 36) and specificity was 89 percent (95% CI, 88 to 91); PPV and NPV were 76 (95% CI, 55 to 89) and 79 (95% CI, 58 to 87), respectively.

Key Question 2. What is the direct evidence on the benefits and harms of screening women for GDM to reduce maternal, fetal, and infant morbidity and mortality?

Description of Included Studies

Two studies met the inclusion criteria for Key Question 2.[130,131] Both studies compared outcomes for women who underwent screening or diagnostic testing for GDM with women who were not screened or tested. The studies are described in Appendix D. The studies were published in 2004[130] and 1996.[131] The methods and outcomes differed between the studies, therefore no results were pooled.

Methodological Quality of Included Studies

The studies were of high and moderate methodological quality with 7 and 6 of a maximum of 9 points, respectively.[130,131] The studies scored well for selection of the non-exposed cohort (same as exposed cohort), ascertainment of exposure and outcome, and adequacy of followup in terms of duration and attrition. Neither study controlled for potential confounding variables. Solomon et al., included a select population (i.e., nurses participating in a longitudinal study) that may not be representative of the general target population of this review.

Key Points

Only two retrospective cohort studies were relevant to Key Question 2. There were no RCTs available to answer questions about screening. Based on the small number of studies and sample sizes, the impact of screening women for GDM on health outcomes is inconclusive.

Detailed Synthesis

One retrospective cohort study examined 1,000 women receiving antenatal care and delivering at a single center in Thailand between October 2001 and December 2002.[130] Women who presented with specific risk factors underwent screening with OGCT (n = 411), and subsequent OGTT if positive on the OGCT (n = 164). Among those screened, 29 cases of GDM were identified (7 percent of the screened group; 3 percent of the total population). Among those who did not undergo screening, 40 women at high risk for GDM were missed (4 percent) and there were two cases of pregestational DM (0.2 percent). High risk was determined based on a list of risk factors, the most commonly observed were age ≥ 30 years (53 percent of the 40 patients) and family history of type 2 diabetes mellitus (43 percent of the 40 patients). Appendix D lists the obstetric complications that were reported in decreasing frequency. Overall there were significantly more complications in the screened group (64/411 versus 63/589). The only individual obstetric complication that was different between groups was pregnancy-induced hypertension with significantly more cases in the screened group. The screened group was significantly older and had a higher average BMI than the group not screened. The pregnancy outcomes are listed in Appendix D. The only significant difference was in the incidence of cesarean deliveries which was greater in the screened group. The authors concluded that selective OGCT screening was highly effective in detecting GDM; however, the impact on outcomes was inconclusive due to small numbers. No information was provided on how women who screened positive were treated.

The second study involved a survey of a subset of participants in a large prospective cohort study involving 116,678 nurses age 25-42 years (the Nurses' Health Study II).[131] Surveys were sent to 422 women who reported a first diagnosis of GDM between 1989 and 1991, as well as a sample of 100 women who reported a pregnancy but no diagnosis of GDM. The intent of the study was to determine the frequency of screening for GDM and the extent to which diagnosis is based on NDDG criteria. Only one outcome was reported that was relevant to this Key Question: the incidence of macrosomia (infant weight > = 4.3 kg) was the same in the screened and unscreened groups (7 percent each group). These results pertained to 93 eligible women who reported a pregnancy and no diagnosis of GDM, 77 of whom reported having a 1-h 50-g OGCT. No information was provided on how women who screened positive were treated. No relevant outcomes were reported for the group of women who reported a pregnancy and first diagnosis of GDM.

Key Question 3. In the absence of treatment, how do health outcomes of mothers who meet various criteria for GDM and their offspring compare to those who do not?

Description of Included Studies

Thirty-eight studies met the inclusion criteria for Key Question 3.[3,54,67,78-94,102,103,106,132-137,142,145-147,149,150,152,154,155] The studies are described in Appendix D. Studies provided data for untreated women who met criteria for GDM, showed differing levels of glucose tolerance, or had no GDM. Most included studies were prospective or retrospective cohort studies published between 1995 and 2011 (median year 2004). Two studies were long-term followup studies of RCTs; however, only data from the untreated patients were included in the results for this Key Question.[54,142] These studies had associated publications providing more detailed break-down of groups and outcomes.[160,163] Fourteen studies were conducted in the U.S.,[54,78,81,88-91,132,135,136,146,150,152] 10 in Europe,[80,86,87,93,102,106,133,145,149,154] 2 in Canada,[83,142] 2 in Australia,[3,85] and 11 from other countries[67,79,82,84,92,94,103,134,137,147,155] (including Japan, Saudi Arabia, Turkey, Iran, China, and Taiwan). Populations analyzed in North American studies involved diverse ethnicities representative of the respective populations; studies from Europe or elsewhere most often included women of ethnic descent from the country of study origin. In one case, women analyzed were at risk for GDM;[149] this study has been noted as potentially unrepresentative of all women eligible for screening.

We grouped studies according to the diagnostic criteria used; these included CC, NDDG, WHO, and IADPSG. CC values were endorsed by the ADA 2000-2010 as well as the 4th and 5th IWC on Gestational Diabetes. Most studies employing NDDG criteria provided comparison groups of women diagnosed with CC criteria. In most cases, the NDDG GDM group received treatment for GDM as it is commonly considered unethical in North America to not treat these women; therefore, these groups were not included in the results for this Key Question. One study compared unrecognized cases of NDDG GDM with a patient group with no GDM; the unrecognized cases were sixteen women diagnosed postpartum and therefore did not receive any treatment.[152] CC groups were included; therefore, data from studies employing NDDG criteria with CC comparison groups, CC criteria, ADA, or 4th – 5th IWC criteria were included in the results. Table 1 provides an overview of these criteria.

Seventeen studies employed NDDG criteria (with treated groups excluded from this analysis), CC criteria, ADA, or 4th-5th IWC criteria with comparable groups. Groups included GDM diagnosed by CC criteria, no GDM by any criteria (normal), impaired glucose tolerance (IGT) defined as one abnormal glucose value (OAV), and false positive (positive OGCT, negative OGTT). Two studies had unique group selections and are described in the text below.

Six studies utilized NDDG criteria exclusively. Four of these presented consistent groups for analysis: normal (no GDM by any criteria) and false positive. One study retrospectively identified women with unrecognized GDM by NDDG criteria and compared this group with woman with normal glucose tolerance.

Eight studies presented data according to WHO criteria, four of which provided comparable groups. WHO criteria proved a significant challenge due to variability by year, studies providing insufficient groupings for comparison, and treatment of most IGT or OAV groups. One of the two included studies provided data for women diagnosed with IGT at 8.0-8.9 mmol/L (untreated) and the other provided a similar IGT diagnosis at 7.8-8.9 mmol/L, both at two hours post 75 g load. Studies were pooled for analysis as they were deemed to be sufficiently similar. One study

compared WHO GDM (untreated) with no GDM, and was included in the analysis for macrosomia.[84] Three studies comparing differing levels of WHO criteria were excluded from pooled analysis because they did not have comparable groups with other included studies.[134,137,147]

Three studies utilized IADPSG criteria for diagnosis and provided comparable groups for pooled analysis.[78,79,93]

Methodological Quality of Included Studies

The methodological quality of the included studies is described in Appendix C3. Quality was analyzed using the Newcastle-Ottawa Scale (NOS) with a possible total of 9 stars. The median quality score was 9 stars, with two studies receiving a score of 6/9, nine studies a score of 7/9, seven studies a score of 8/9, and twenty a score of 9/9. Studies receiving lower scores on the NOS most often did not control for potential confounding (e.g., due to BMI, age, race), and/or had an important proportion of patients lost to followup. Overall, the majority of studies were considered good quality (36 of 38, 95 percent).

Key Points

- Thirty-eight studies provided data for this question that sought to examine health outcomes for women who meet various criteria for GDM and do not receive treatment. The majority of data came from cohort studies or the untreated groups from randomized trials.
- A wide variety of diagnostic criteria and thresholds were compared across the studies. The most common groups reported and compared were GDM diagnosed by CC criteria, no GDM by any criteria (normal), impaired glucose tolerance defined as OAV, and false positive (positive OGCT, negative OGTT). The following criteria were used: CC (19 studies), NDDG (6 studies), WHO (8 studies), and IADPSG (3 studies).

Maternal Outcomes

- A methodologically strong study showed a continuous positive relationship between increasing glucose levels and the incidence of primary cesarean section. This study also found significantly fewer cases of preeclampsia and cesarean section among women without GDM compared with those meeting IADPSG criteria.
- For preeclampsia, significant differences were found for CC versus patients with no GDM (3 studies) with fewer cases among the patients with no GDM, and for CC GDM versus false-positive groups (2 studies) with fewer cases among the false positives. The strength of evidence for these comparisons was low. No differences were found for NDDG false positive versus no GDM (2 studies), NDDG 1 abnormal OGTT versus no GDM (1 study), and IGT WHO versus no GDM (3 studies); the strength of evidence for these findings was insufficient.
- For maternal hypertension, significant differences were found for eight of 16 comparisons; five of these comparisons were based on single studies. Patient groups with no GDM showed lower incidence of maternal hypertension when compared with CC GDM, CC false positives, CC 1 abnormal OGTT, IADPSG impaired fasting glucose (IFG), IADPSG double impaired glucose tolerance (IGT-2), and IADPSG IGT IFG. Other comparisons showing significant differences were CC GDM versus false positives

(lower incidence for false positives), IADPSG IGT versus IGT IFG (lower incidence for IGT), and IADPSG IFG versus IGT IFG (lower incidence for IFG).

- There were 21 comparisons for cesarean section with nine significant differences. Patient groups with no GDM showed fewer cesarean sections when compared with CC GDM (9 studies), CC 1 abnormal OGTT (4 studies), CC false positives (5 studies), NDDG false positives (4 studies), NDDG 1 abnormal OGTT (1 study), and WHO IGT (4 studies). Four studies compared CC GDM versus false positives and showed lower incidence for the false positives. Single studies compared IADPSG IFG and IADPSG IGT IFG versus no GDM, respectively, and both showed fewer cases for the patient groups with no GDM.
- Based on single studies, no differences were observed for maternal birth trauma for CC GDM versus no GDM , CC GDM versus false positives, NDDG GDM (unrecognized) versus no GDM.
- For maternal weight gain, significant differences were found for three of 12 comparisons: IADPSG IGT versus no GDM (favored IGT), IADPSG IFG versus no GDM (favored IFG), IADPSG IGT-2 versus no GDM (favored IGT-2). All comparisons were based on single studies and strength of evidence was considered insufficient.
- For maternal mortality/morbidity, single studies compared CC GDM versus no GDM, CC 1 abnormal OGTT versus no GDM, IADPSG GDM versus no GDM. No differences were found except for the latter comparison that showed lower mortality/morbidity for the patient groups with no GDM.
- No studies provided data on long-term maternal outcomes, such as type 2 diabetes mellitus, obesity and hypertension.

Fetal/Neonatal/Child Outcomes

- Two methodologically strong studies showed a continuous positive relationship between increasing glucose levels and the incidence of macrosomia. One of these studies also showed significantly fewer cases of shoulder dystocia and/or birth injury, clinical neonatal hypoglycemia, and hyperbilirubinemia among women without GDM compared with women meeting IADPSG criteria.
- The most commonly reported outcome was macrosomia >4,000 g. Eleven comparisons were made of which six showed a significant difference. Fewer cases were observed among patient groups with no GDM compared with CC GDM (10 studies), CC 1 abnormal OGTT (7 studies), NDDG GDM (unrecognized) (1 study), NDDG false-positives (4 studies), and WHO IGT (1 study). Fewer cases were found for women with false-positive results compared with CC GDM (5 studies). The strength of evidence for these findings was low to insufficient.
- Data for macrosomia >4,500 g were available for four comparisons and showed significant differences in two cases: patient groups with no GDM had fewer cases compared with women with CC GDM and with unrecognized NDDG GDM. The strength of evidence for these findings was low and was insufficient, respectively.
- For shoulder dystocia, significant differences were found for 7 of 17 comparisons; all but 1 comparison was based on single studies (insufficient strength of evidence). Patient groups with no GDM showed lower incidence of shoulder dystocia when compared with CC GDM (5 studies; low strength of evidence), NDDG GDM (unrecognized), NDDG false positive, WHO IGT, IADPSG IFG, and IADPSG IGT IFG. The other significant

difference showed lower incidence among the false-positive group compared with CC 1 abnormal OGTT.

- For fetal birth trauma/injury, four studies compared CC GDM, NDDG GDM, and WHO IGT with no GDM. No differences were observed except for NDDG GDM which favored the patient group with no GDM. Strength of evidence was insufficient for all comparisons.
- Only one difference was found for neonatal hypoglycemia with fewer cases among patient groups with no GDM compared with those meeting CC criteria. No differences were found for other comparisons, including CC GDM versus 1 abnormal OGTT (1 study), CC 1 abnormal OGTT versus no GDM (4 studies), NDDG GDM versus no GDM (1 study), NDDG false positive versus no GDM (1 study), NDDG 1 abnormal OGTT versus no GDM (1 study), and WHO IGT versus no GDM (3 studies). Strength of evidence was insufficient for all comparisons.
- There were 16 comparisons for hyperbilirubinemia; the majority were based on single studies. Three comparisons showed significant differences between groups: patient groups with no GDM had fewer cases compared with CC false positive, IADPSG IGT, and IADPSG IGT-2, respectively.
- No differences were found for fetal morbidity/mortality for any of 8 comparisons which may be attributable to small numbers of events within some comparisons. Most comparisons were based on few studies, except for CC GDM versus no GDM which showed no difference based on 6 studies.
- Based on single studies, significant differences were found in prevalence of childhood obesity for CC GDM versus groups with no GDM (lower prevalence for no GDM) and CC GDM versus false positives (lower prevalence for false positives). No differences, based on single studies, were found for CC GDM versus 1 abnormal OGTT, CC false positive versus no GDM, CC false positive versus 1 abnormal OGTT, or CC 1 abnormal OGTT versus no GDM. No other studies provided data on long-term outcomes, including type 2 diabetes mellitus and transgenerational GDM.

Detailed Synthesis

Overview

Detailed results are described by outcome in the sections that follow. We first describe the maternal outcomes, followed by fetal/neonatal/child outcomes. We present meta-graphs when two or more studies were pooled. These are displayed after the description of results for each outcome. A detailed table of results and a table summarizing the strength of evidence are presented at the end of each of the maternal and fetal/neonatal/child sections (Table 12 and

Table 13; Table 14 and Table 15, respectively). The results reported below are based on unadjusted data from the relevant studies. We have reported adjusted results, where available from relevant studies, in Appendix G. In the majority of cases, the adjusted results would not have changed the pooled estimates or overall conclusions. Six studies met inclusion criteria and provided relevant outcomes but were not comparable with other studies and are described here.[3,91,134,137,147]

In 1995, Sacks et al. published a prospective cohort study of 3,505 unselected pregnant women; the authors sought to determine glucose threshold distributions for the 2 hr, 75 g OGTT, and to define the relationship between glucose intolerance values and neonatal macrosomia. The

methodological quality of the study was good receiving a score of 8/9 points. Study participants were not analyzed by groups, rather regression analyses were conducted to identify a threshold level that predicted greater risk for macrosomia. The study did not identify a specific threshold for fasting or 1-2 hour levels that could discriminate between women who were more likely to have infants with macrosomia. Moreover, across all thresholds the ability to predict macrosomia was relatively consistent.

The HAPO (Hyperglycemia and Adverse Pregnancy Outcomes) study, published in 2008, examined the effect of less severe hyperglycemia on pregnancy outcomes; therefore, all groups fell below the common diagnostic thresholds for GDM. The study involved 23,316 pregnant women from 15 centers in nine countries. The methodological quality was good with a score of 9/9 points. Women were tested employing the 75 g OGTT at 24-32 weeks. Fasting plasma glucose values were divided into seven categories: ≥100 mg/dL (5.6 mmol/L), 95-99 (5.3-5.5), 90-94 (5.0-5.2), 85-89 (4.8-4.9), and <85. The last category (<85 mg/dL) was further subdivided into three levels: <75 mg/dL (4.2 mmol/L), 75-59 (4.2-4.4), and 80-84 (4.5-4.7). The study found a continuous positive association with increasing glucose levels and macrosomia (or birthweight >90th percentile), primary cesarean section, neonatal hypoglycemia, and cord-blood serum c-peptide >90th percentile. The associations were strongest for macrosomia and blood serum c-peptide levels; moreover, associations for neonatal hypoglycemia were not consistently significant. In unadjusted analyses, preeclampsia, cesarean delivery, shoulder dystocia and/or birth injury, clinical neonatal hypoglycemia, and hyperbilirubinemia were statistically significantly less frequent for women without GDM compared with those with GDM based on the IADPSG criteria (data from Appendix, Table B available at care.diabetesjournals.org/cgi/content/full/dc09-1848/DC1). The study did not identify a clear glucose threshold for increased risk in clinically important outcomes.[24]

Two studies[134,147] conducted in China utilized 1980 WHO criteria on a 2 hr OGTT but did not provide similar groups for comparison. One retrospective cohort study published in 2003 involving 2,149 women compared six glucose values: <6.0 mmol/L, 6.0-6.9, 7.0-7.9, 8.0-8.9, 9.0-10.9, and ≥11.0.[147] The latter 3 groups were treated for GDM; the former were untreated. There was no significant difference between groups in the incidence of macrosomia (≥4,000 g) or cesarean deliveries. The methodological quality of the study was good with 8/9 points. The second study published in 2001 was prospective and involved 487 women. The study compared a control group, an "at risk" but normal OGTT group, and a treated GDM group.[134] There were no significant differences between groups in preeclampsia or birthweight. There were significantly more cesarean deliveries in the normal OGTT compared with the control group although the comparison did not control for age and BMI (women in the normal OGTT group were older and more obese). The methodological quality was fair scoring 6/9 points.

One study[137] conducted in Malaysia used 1999 WHO criteria on a 2 hr OGTT in conjunction with a 50 g OGCT. As WHO criteria rarely utilize an OGCT, this study did not provide comparable groups for pooled analysis as they were based upon OGCT test results. The study found significantly more cases of cesarean delivery, postpartum hemorrhage, and macrosomia (>4,000 g) among OGCT-positive versus OGCT-negative women.

A study conducted in Turkey between 2003 and 2009 employed CC criteria on a 50 g OGCT as well as a 3 hr, 100 g OGTT.[94] Groups were determined according to abnormal fasting, 1 hr, 2 hr, and 3 hr glucose values, which did not provide comparison to included studies. The study did not find a significant difference between groups in mean neonatal birthweight. There were

significantly more cases of macrosomia (>4,000 g) among women with increased serum glucose at 2 hours.

Maternal Outcomes

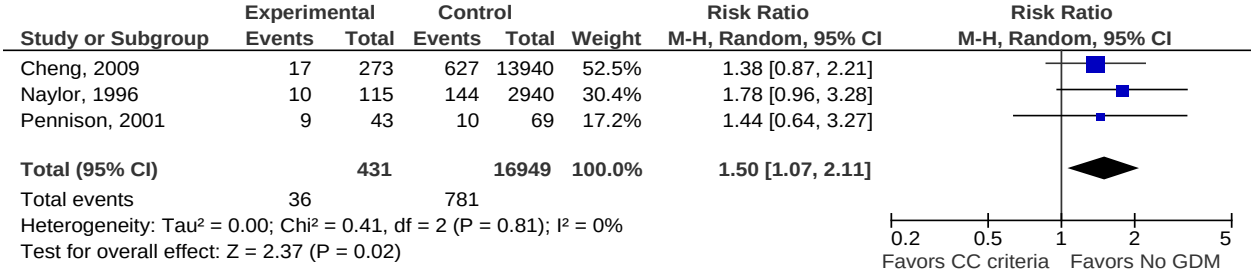
Short Term

A summary of the evidence for short-term maternal outcomes is provided in Table 12. A summary of the strength of evidence is in Table 13. The sections that follow describe the results by outcome.

Preeclampsia

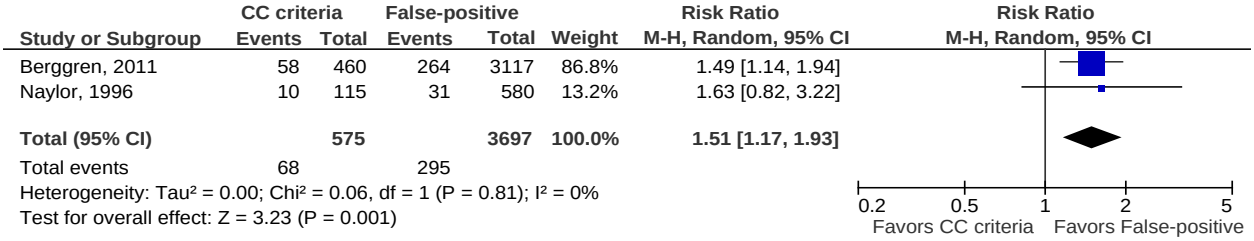
Ten studies presented data on preeclampsia (Table 12).[81,82,88-90,103,133,149,155,160] Definitions of preeclampsia were only reported in two of the ten studies, and the definitions differed. Three studies compared women who met CC criteria for GDM with women who had no GDM and found a significant difference with fewer cases among the no GDM group (Figure 13).[81,89,160] Two studies compared women who met CC criteria for GDM with women who were false positive and demonstrated a significant difference with fewer cases in the false-positive group (Figure 14).[90,160] The strength of evidence for these two comparisons was low. The following three comparisons showed no differences between groups: 1 abnormal OGTT by NDDG versus no GDM (1 study),[103] false positive NDDG versus no GDM (2 studies, Figure 15),[82,88] and IGT by WHO criteria versus no GDM (3 studies, Figure 16).[133,149,155] The strength of evidence for these three comparisons was insufficient.

Figure 13. CC GDM versus no GDM: preeclampsia



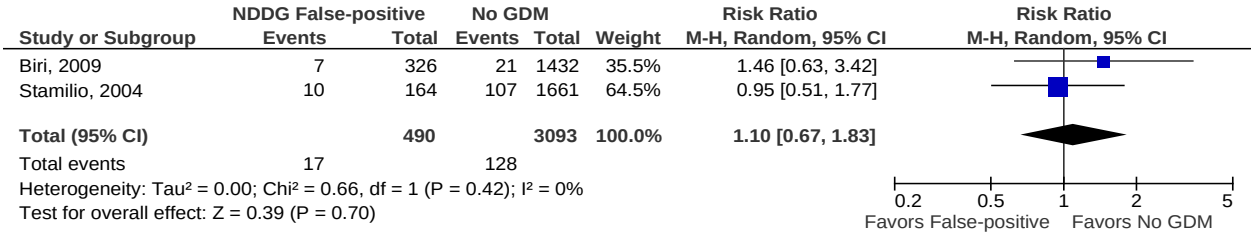
CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus, M-H = Mantel-Haenszel

Figure 14. CC GDM versus false positive: preeclampsia



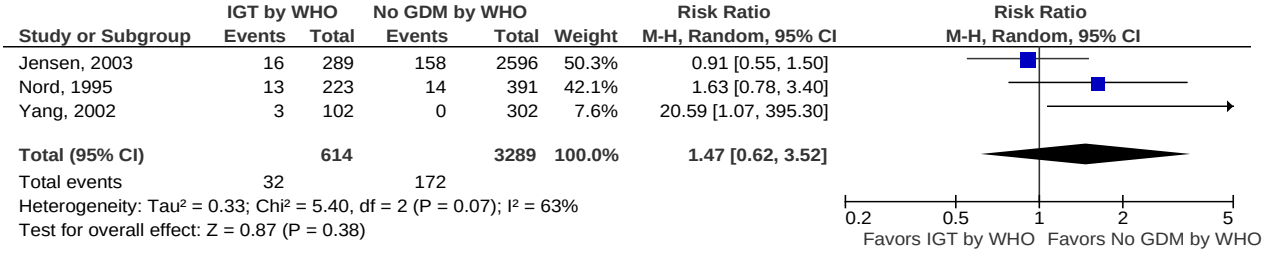
CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus, M-H = Mantel-Haenszel

Figure 15. NDDG false positive versus no GDM: preeclampsia



CI = confidence interval; GDM = gestational diabetes mellitus, NDDG = National Diabetes Data Group; M-H = Mantel-Haenszel

Figure 16. WHO impaired glucose tolerance versus no GDM: preeclampsia

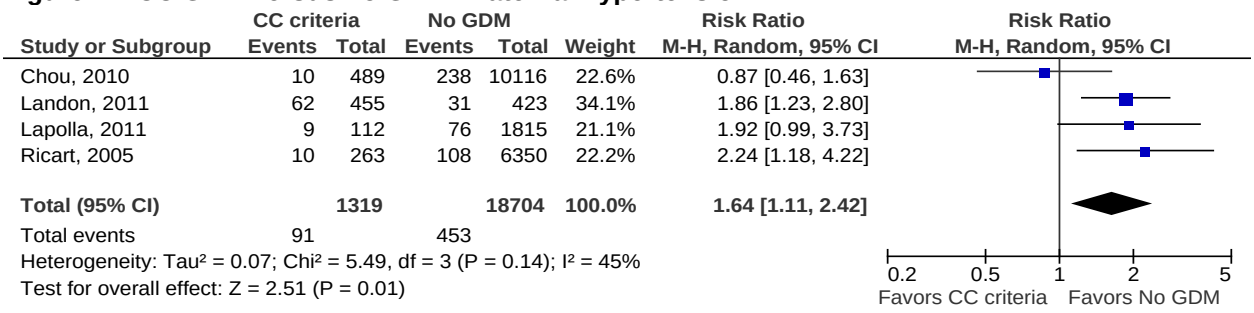


CI = confidence interval; GDM = gestational diabetes mellitus, IGT = impaired glucose tolerance; M-H = Mantel-Haenszel; WHO = World Health Organization

Maternal Hypertension

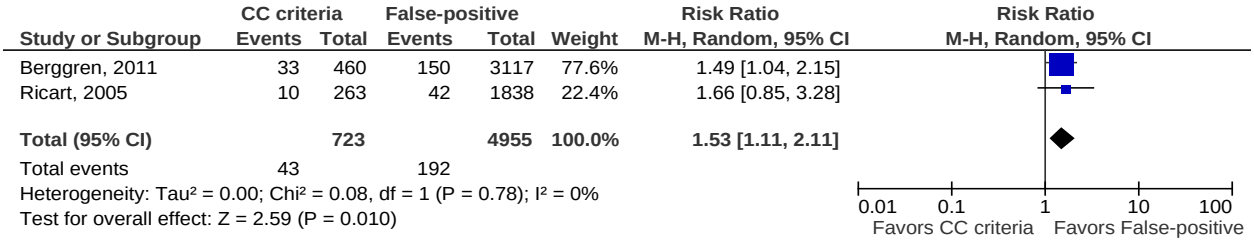
Nine studies presented data on maternal hypertension (Table 12).[78,80,90,92,93,102,106,133,163] Four studies compared women who met CC criteria for GDM with women without GDM and showed significantly fewer cases in the no GDM group (Figure 17).[92,93,102,163] Two studies comparing women who met CC criteria for GDM with women who were false positive showed a significant difference with fewer cases in the false-positive group (Figure 18).[90,102] Two studies compared one abnormal OGTT by CC criteria with no GDM and showed a significant difference with fewer cases in the group with no GDM (Figure 19).[80,106] No differences were found for the following comparisons: CC false positive versus no GDM (1 study),[102] WHO IGT versus no GDM (1 study),[133] and IADPSG GDM versus no GDM (1 study).[93] A single study of IADPSG criteria[78] made comparisons across six different groups and found significant differences for: IADPSG IFG versus no GDM, IADPSG double impaired glucose tolerance (IGT-2) versus no GDM, IADPSG IGT IFG versus no GDM (all favoring no GDM); IADPSG IGT versus IGT IFG (favoring IGT); and IADPSG IFG versus IGT IFG (favoring IFG).

Figure 17. CC GDM versus no GDM: maternal hypertension



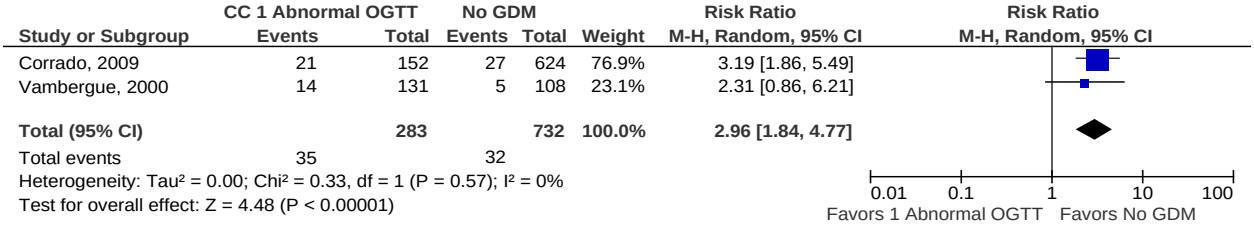
CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus, M-H = Mantel-Haenszel

Figure 18. CC GDM versus false positive: maternal hypertension



CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus, M-H = Mantel-Haenszel

Figure 19. CC 1 Abnormal OGTT versus no GDM: maternal hypertension



CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus, M-H = Mantel-Haenszel ; OGTT = oral glucose tolerance test

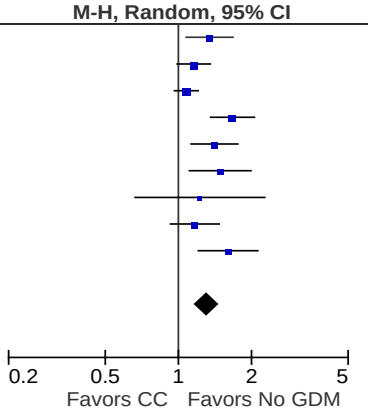
Cesarean Delivery

Twenty-six studies presented data for cesarean delivery (Table 12).[67,78,80,81,83,85-90,92,93,102,103,132,133,135,145,146,149,150,152,154,155,160] Nine studies compared CC GDM with no GDM and found a significant difference with fewer cases for the no GDM group (Figure 20).[81,86,89,92,93,102,146,150,160] Four studies compared CC GDM with false-positive results and showed significantly fewer cases in the false-positive group (Figure 21).[90,102,150,160] Four studies compared CC 1 abnormal OGTT versus no GDM and found fewer cases in the group with no GDM (Figure 22).[80,86,106,135] Five studies compared CC false positives with no GDM and found fewer events among patient groups with no GDM (Figure 23).[87,102,145,150,160] One study compared NDDG with 1 abnormal OGTT with women without GDM and found fewer events for the no GDM group.[103] Four studies comparing NDDG false positives versus no GDM showed a significant difference with fewer events for the no GDM group (Figure 24).[67,88,132,152] Four studies compared WHO impaired glucose tolerance with no GDM, a significant difference was found in favor of the no GDM group (Figure 25).[133,149,154,155] One study compared IADPSG IFG versus no GDM, and the same study compared IADPSG IGT IFG versus no GDM with both showing significant differences with fewer cases in the no GDM group.[78] There were no differences between groups for the remaining comparisons (Table 12; Figure 26).

Figure 20. CC GDM versus no GDM: cesarean delivery

Study or Subgroup	CC Events	CC Total	No GDM Events	No GDM Total	Weight	Risk Ratio M-H, Random, 95% CI
Cheng, 2009	62	273	2356	13940	11.7%	1.34 [1.08, 1.68]
Chico, 2005	122	422	1442	5767	14.7%	1.16 [0.99, 1.35]
Chou, 2010	196	489	3761	10116	16.8%	1.08 [0.96, 1.20]
Langer, 2005	132	555	158	1110	12.3%	1.67 [1.36, 2.06]
Lapolla, 2011	49	112	564	1815	11.7%	1.41 [1.13, 1.76]
Naylor, 1996	34	115	585	2940	9.0%	1.49 [1.11, 1.99]
Pennison, 2001	13	43	17	69	3.2%	1.23 [0.66, 2.27]
Ricart, 2005	59	263	1219	6350	11.3%	1.17 [0.93, 1.47]
Schwartz, 1999	38	154	1110	7207	9.4%	1.60 [1.21, 2.12]
Total (95% CI)		**2426**		**49314**	**100.0%**	**1.32 [1.17, 1.48]**
Total events	705		11212			

Heterogeneity: Tau² = 0.02; Chi² = 21.49, df = 8 (P = 0.006); I² = 63%
Test for overall effect: Z = 4.54 (P < 0.00001)

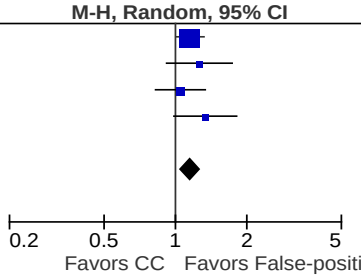


CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus, M-H = Mantel-Haenszel

Figure 21. CC GDM versus false positive: cesarean delivery

Study or Subgroup	CC Events	CC Total	False-positive Events	False-positive Total	Weight	Risk Ratio M-H, Random, 95% CI
Berggren, 2011	160	460	942	3117	58.8%	1.15 [1.00, 1.32]
Naylor, 1996	34	115	136	580	10.7%	1.26 [0.92, 1.73]
Ricart, 2005	59	263	393	1838	18.7%	1.05 [0.82, 1.34]
Schwartz, 1999	38	154	197	1066	11.8%	1.34 [0.99, 1.81]
Total (95% CI)		**992**		**6601**	**100.0%**	**1.16 [1.05, 1.29]**
Total events	291		1668			

Heterogeneity: Tau² = 0.00; Chi² = 1.77, df = 3 (P = 0.62); I² = 0%
Test for overall effect: Z = 2.83 (P = 0.005)

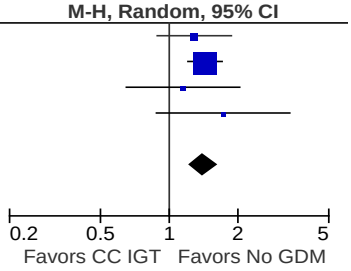


CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus, M-H = Mantel-Haenszel

Figure 22. CC, 1 abnormal OGTT versus no GDM: cesarean delivery

Study or Subgroup	1 Abnormal OGTT Events	1 Abnormal OGTT Total	No GDM Events	No GDM Total	Weight	Risk Ratio M-H, Random, 95% CI
Chico, 2005	19	59	1442	5767	15.5%	1.29 [0.89, 1.87]
Corrado, 2009	85	152	243	624	73.1%	1.44 [1.21, 1.71]
Rust, 1996	14	78	32	205	6.6%	1.15 [0.65, 2.04]
Vambergue, 2000	23	131	11	108	4.8%	1.72 [0.88, 3.37]
Total (95% CI)		**420**		**6704**	**100.0%**	**1.40 [1.21, 1.63]**
Total events	141		1728			

Heterogeneity: Tau² = 0.00; Chi² = 1.10, df = 3 (P = 0.78); I² = 0%
Test for overall effect: Z = 4.52 (P < 0.00001)

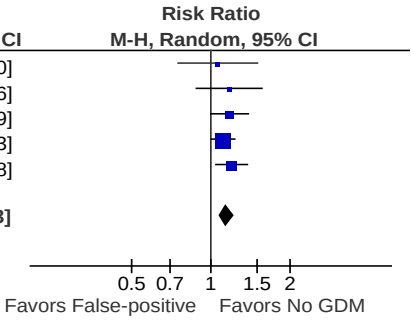


CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus, IGT = impaired glucose tolerance; M-H = Mantel-Haenszel ; OGTT = oral glucose tolerance test

Figure 23. CC false positive versus no GDM: cesarean delivery

Study or Subgroup	False-positive Events	Total	No GDM Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Bo, 2004	103	315	28	91	4.0%	1.06 [0.75, 1.50]
Lapolla, 2007	45	128	100	334	5.8%	1.17 [0.88, 1.56]
Naylor, 1996	136	580	585	2940	17.8%	1.18 [1.00, 1.39]
Ricart, 2005	393	1838	1219	6350	46.9%	1.11 [1.01, 1.23]
Schwartz, 1999	197	1066	1110	7207	25.5%	1.20 [1.05, 1.38]
Total (95% CI)		**3927**		**16922**	**100.0%**	**1.15 [1.07, 1.23]**
Total events	874		3042			

Heterogeneity: Tau² = 0.00; Chi² = 1.05, df = 4 (P = 0.90); I² = 0%
Test for overall effect: Z = 3.91 (P < 0.0001)

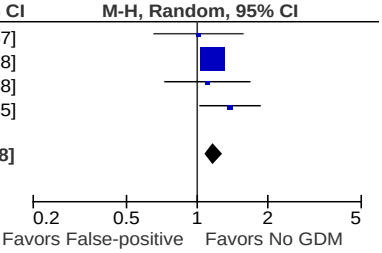


CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus, M-H = Mantel-Haenszel

Figure 24. NDDG false-positive versus no GDM: cesarean delivery

Study or Subgroup	False-positive by NDDG Events	Total	No GDM Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Ardawi, 2000	24	187	67	529	3.9%	1.01 [0.66, 1.57]
Hillier, 2007	208	326	785	1432	83.2%	1.16 [1.06, 1.28]
Retnakaran, 2008	44	128	23	74	4.3%	1.11 [0.73, 1.68]
Stamilio, 2004	39	164	286	1661	8.6%	1.38 [1.03, 1.85]
Total (95% CI)		**805**		**3696**	**100.0%**	**1.17 [1.08, 1.28]**
Total events	315		1161			

Heterogeneity: Tau² = 0.00; Chi² = 1.73, df = 3 (P = 0.63); I² = 0%
Test for overall effect: Z = 3.62 (P = 0.0003)

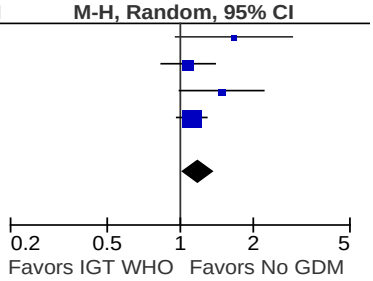


CI = confidence interval; GDM = gestational diabetes mellitus, M-H = Mantel-Haenszel; NDDG = National Diabetes Data Group

Figure 25. WHO impaired glucose tolerance versus no GDM: cesarean delivery

Study or Subgroup	IGT WHO Events	Total	No GDM Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Aberg, 2001	12	131	249	4526	7.0%	1.67 [0.96, 2.89]
Jensen, 2003	54	289	450	2596	26.4%	1.08 [0.84, 1.39]
Nord, 1995	38	223	45	391	12.7%	1.48 [0.99, 2.21]
Yang, 2002	75	102	199	302	53.9%	1.12 [0.97, 1.29]
Total (95% CI)		**745**		**7815**	**100.0%**	**1.18 [1.01, 1.37]**
Total events	179		943			

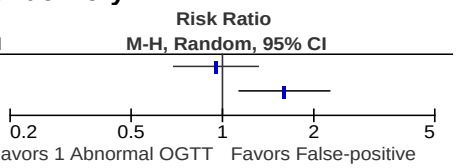
Heterogeneity: Tau² = 0.01; Chi² = 3.86, df = 3 (P = 0.28); I² = 22%
Test for overall effect: Z = 2.12 (P = 0.03)



CI = confidence interval; GDM = gestational diabetes mellitus, IGT = impaired glucose tolerance; M-H = Mantel-Haenszel; WHO = World Health Organization

Figure 26. CC, 1 abnormal OGTT versus false positive: cesarean delivery

Study or Subgroup	1 Abnormal OGTT Events	Total	False-positive Events	Total	Risk Ratio M-H, Random, 95% CI
Kwik, 2007	46	156	61	197	0.95 [0.69, 1.31]
Lapolla, 2007	27	48	45	128	1.60 [1.14, 2.25]



CC = Carpenter-Coustan; CI = confidence interval; M-H = Mantel-Haenszel; OGTT = Oral glucose tolerance test

Birth Trauma

Three studies presented data for maternal birth trauma (Table 12).[81,90,152] Two studies employed CC GDM and compared with no GDM and a false-positive group, respectively.[81,90] In both studies birth trauma was defined as third or fourth degree perineal laceration. Neither study found a significant difference between groups. One study compared unrecognized NDDG GDM with no GDM and showed no difference in rectal injury between groups.[152]

Weight Gain

Three studies presented data for maternal weight gain (Table 12).[78,135,155] One study compared 1 abnormal glucose tolerance value by CC criteria with no GDM and found no difference between groups.[135] One study compared impaired glucose tolerance by WHO criteria with no GDM; no significant difference was found between groups.[155] One study compared varying degrees of glucose intolerance by IADPSG criteria.[78] Significantly less weight gain was found in the IGT, IFG, and IGT-2 groups in comparison with no GDM. No significant differences were noted between any other IADPSG glucose tolerance groups.

Maternal Morbidity/Mortality

Two studies presented data for maternal mortality or morbidity (Table 12).[93,135] One study compared CC GDM as well as IADPSG GDM with no GDM.[93] No significant difference was found between the CC and no GDM groups, while a significant difference favoring no GDM was found in comparison with the IADPSG group. One study compared one abnormal glucose value by CC criteria with no GDM, with no significant difference noted between groups.[135]

Long Term

No studies provided data on long-term maternal outcomes, such as type 2 diabetes mellitus, obesity and hypertension.

Table 12. Evidence summary table: maternal outcomes

Outcome	Comparison	Studies	Participants	Effect Estimate*	I^2	Favors¶
Preeclampsia	CC GDM vs. no GDM	3	17,380	1.50 [1.07, 2.11]	0%	No GDM
	CC GDM vs. false positive	2	4,272	1.51 [1.17, 1.93]	0%	False positive
	NDDG false positive vs. no GDM	2	3,583	1.10 [0.67,1.83]	0%	-
	NDDG, 1 abnormal OGTT vs. no GDM	1	699	1.33 [0.48, 3.65]	NA	-
	WHO IGT vs. no GDM	3	3,903	1.47 [0.62, 3.52]	63 %	-
Maternal hypertension	CC GDM vs. no GDM	4	20,023	1.64 [1.11, 2.42]	45 %	No GDM
	CC GDM vs. false positive	2	5,678	1.53 [1.11, 2.11]	0%	False positive
	CC 1 abnormal OGTT vs. no GDM	2	1,015	2.96 [1.84, 4.77]	0%	No GDM
	CC false positive vs. no GDM	1	8,188	1.35 [0.94, 1.94]	NA	-
	IGT WHO vs. no GDM	1	2,885	0.91 [0.55, 1.50]	NA	-
	IADPSG GDM vs. no GDM	1	1,927	1.92 [0.99, 3.73]	NA	-
	IADPSG IGT vs. no GDM	1	7,411	1.32 [0.96, 1.82]	NA	-

Table 12. Evidence summary table: maternal outcomes (continued)

Outcome	Comparison	Studies	Participants	Effect Estimate*	I^2	Favors¶
Maternal Hypertension (continued)	IADPSG IFG vs. no GDM	1	7,906	1.46 [1.18, 1.80]	NA	No GDM
	IADPSG IGT-2 vs. no GDM	1	7,103	1.90 [1.09, 3.31]	NA	No GDM
	IADPSG IGT IFG vs. no GDM	1	7,351	2.03 [1.54, 2.69]	NA	No GDM
	IADPSG IGT vs. IFG	1	1,277	0.91 [0.63, 1.31]	NA	-
	IADPSG IGT vs. IGT-2	1	474	0.69 [0.37, 1.31]	NA	-
	IADPSG IGT vs. IGT IFG	1	722	0.65 [0.43, 0.98]	NA	IGT
	IADPSG IFG vs. IGT-2	1	969	0.77 [0.43, 1.37]	NA	-
	IADPSG IFG vs. IGT IFG	1	1,217	0.72 [0.51, 0.99]	NA	IFG
	IADPSG IGT-2 vs. IGT IFG	1	414	0.93 [0.51, 1.72]	NA	-
Cesarean delivery	CC GDM vs. no GDM	9	51,740	1.34 [1.17, 1.48]	63%	No GDM
	CC GDM vs. false positive	4	7,593	1.16 [1.05, 1.29]	0%	False positive
	CC GDM vs. 1 abnormal OGTT	1	481	0.90 [0.60, 1.34]	NA	-
	CC 1 abnormal OGTT vs. No GDM	4	7,124	1.40 [1.21, 1.63]	0%	No GDM
	CC false positive vs. no GDM	5	20,849	1.15 [1.07, 1.23]	0%	No GDM
	CC 1 abnormal OGTT vs. false positive	2	529	Results not pooled due to substantial heterogeneity.	79%	-
	NDDG GDM (unrecognized) vs. no GDM	1	80	1.60 [0.58, 4.45]	NA	-
	NDDG, 1 abnormal OGTT vs. no GDM	1	699	1.69 [1.04,2.75]	NA	No GDM
	NDDG false positive vs. no GDM	4	4,501	1.17 [1.08, 1.28]	0%	No GDM
	WHO IGT vs. no GDM	4	8,560	1.18 [1.01, 1.37]	22%	No GDM
	IADPSG GDM vs. no GDM	1	1,927	1.92 [0.99, 3.73]	NA	-
	IADPSG IGT vs. no GDM	1	7,411	1.11 [0.89, 1.39]	NA	-
	IADPSG IFG vs. no GDM	1	7,906	1.28 [1.11, 1.47]	NA	No GDM
	IADPSG IGT-2 vs. no GDM	1	7,103	1.58 [0.94, 2.64]	NA	-
	IADPSG IGT IFG vs. no GDM	1	7,351	1.32 [1.06, 1.63]	NA	No GDM
	IADPSG IGT vs. IFG	1	1,277	0.87 [0.68, 1.12]	NA	-

Table 12. Evidence summary table: maternal outcomes (continued)

Outcome	Comparison	Studies	Participants	Effect Estimate*	I^2	Favors[¶]
Cesarean delivery (continued)	IADPSG IGT vs. IGT-2	1	474	0.77 [0.49, 1.21]	NA	-
	IADPSG IGT vs. IGT IFG	1	722	0.85 [0.63, 1.14]	NA	-
	IADPSG IFG vs. IGT-2	1	969	0.88 [0.58, 1.34]	NA	-
	IADPSG IFG vs. IGT IFG	1	1,217	0.97 [0.76, 1.24]	NA	-
	IADPSG IGT-2 vs. IGT IFG	1	414	1.10 [0.70, 1.72]	NA	-
	IADPSG IGT vs. IGT IFG	1	722	0.85 [0.63, 1.14]	NA	-
	IADPSG IFG vs. IGT-2	1	969	0.88 [0.58, 1.34]	NA	-
	IADPSG IFG vs. IGT IFG	1	1,217	0.97 [0.76, 1.24]	NA	-
	IADPSG IGT-2 vs. IGT IFG	1	414	1.10 [0.70, 1.72]	NA	-
Maternal birth trauma	CC GDM vs. no GDM	1	14,213	1.26 [0.90, 1.76]	NA	-
	CC GDM vs. false positive	1	3,577	0.80 [0.47, 1.39]	NA	-
	NDDG GDM (unrecognized) vs. No GDM	1	80	2.00 [0.40, 9.97]	NA	-
Maternal weight gain	CC 1 abnormal OGTT vs. no GDM	1	283	Not calculated[†]	NA	-
	WHO IGT vs. no GDM	1	404	0.00 [-1.41, 1.41]	NA	-
	IADPSG IGT vs. no GDM	1	7,411	-1.90 [-3.37, -0.43][‡]	NA	IGT
	IADPSG IFG vs. no GDM	1	7,906	-1.20 [-2.25, -0.15][‡]	NA	IFG
	IADPSG IGT-2 vs. no GDM	1	7,103	-2.60 [-5.12, -0.08][‡]	NA	IGT-2
	IADPSG IGT IFG vs. no GDM	1	7,351	-1.20 [-2.83, 0.43][‡]	NA	-
	IADPSG IGT vs. IFG	1	1,277	-0.70 [-2.45, 1.05][‡]	NA	-
	IADPSG IGT vs. IGT-2	1	474	0.70 [-2.18, 3.58][‡]	NA	-
	IADPSG IGT vs. IGT IFG	1	722	-0.70 [-2.85, 1.45][‡]	NA	-
	IADPSG IFG vs. IGT-2	1	969	1.40 [-1.29, 4.09][‡]	NA	-
	IADPSG IFG vs. IGT IFG	1	1,217	0.00 [-1.88, 1.88][‡]	NA	-
	IADPSG IGT-2 vs. IGT IFG	1	414	-1.40 [-4.36, 1.56][‡]	NA	-

Table 12. Evidence summary table: maternal outcomes (continued)

Outcome	Comparison	Studies	Participants	Effect Estimate*	I^2	Favors¶
Maternal mortality/ morbidity	CC GDM vs. no GDM	1	1,927	1.53 [0.97, 2.42]	NA	-
	CC 1 abnormal OGTT vs. no GDM	1	283	1.01 [0.37, 2.74]	NA	-
	IADPSG GDM vs. no GDM	1	1,927	1.43 [1.01, 2.04]	NA	No GDM

CC = Carpenter-Coustan; GDM = gestational diabetes mellitus; IFG = impaired fasting glucose; IGT = impaired glucose tolerance; IGT-2 = double impaired glucose tolerance; IADPSG = International Association of Diabetes and Pregnancy Study Groups; NA = not applicable; NDDG = National Diabetes Data Group; OGTT = oral glucose tolerance test; WHO = World Health Organization

*Effect estimates are risk ratios and 95% confidence intervals, except where indicated.

¶Where the result was statistically significant, we have listed the group that had the better outcome (e.g., lower incidence of preeclampsia).

†Study did not report variances but did report no significant difference between groups.

‡Effect estimates are mean differences and 95% confidence intervals.

Table 13. Strength of evidence summary table: maternal outcomes

Outcome	Comparison	Studies	Risk of Bias	Consistency	Directness	Precision	SOE
Preeclampsia	CC GDM vs. no GDM	3	High	Consistent	Direct	Precise	Low
	CC GDM vs. false positive	2	High	Consistent	Direct	Precise	Low
	NDDG false positive vs. no GDM	2	High	Consistent	Direct	Imprecise	Insufficient
	NDDG, 1 abnormal OGTT vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	WHO IGT vs. no GDM	3	High	Consistent	Direct	Imprecise	Insufficient
Maternal weight gain	CC 1 abnormal OGTT vs. no GDM	1	High	Unknown	Direct	Unknown	Insufficient
	WHO IGT vs. no GDM	1	High	Unknown	Direct	Unknown	Insufficient
	IADPSG IGT vs. no GDM	1	High	Unknown	Direct	Precise	Insufficient
	IADPSG IFG vs. no GDM	1	High	Unknown	Direct	Precise	Insufficient
	IADPSG IGT-2 vs. no GDM	1	High	Unknown	Direct	Precise	Insufficient
	IADPSG IGT IFG vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IGT vs. IFG	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IGT vs. IGT-2	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IGT vs. IGT IFG	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IFG vs. IGT-2	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IFG vs. IGT IFG	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IGT-2 vs. IGT IFG	1	High	Unknown	Direct	Imprecise	Insufficient

CC = Carpenter-Coustan; GDM = gestational diabetes mellitus; IFG = impaired fasting glucose; IGT = impaired glucose tolerance; IGT-2 = double impaired glucose tolerance; IADPSG = International Association of Diabetes and Pregnancy Study Groups; NDDG = National Diabetes Data Group; OGTT = oral glucose tolerance test; SOE = strength of evidence; WHO = World Health Organization

Fetal/Neonatal/Child Outcomes

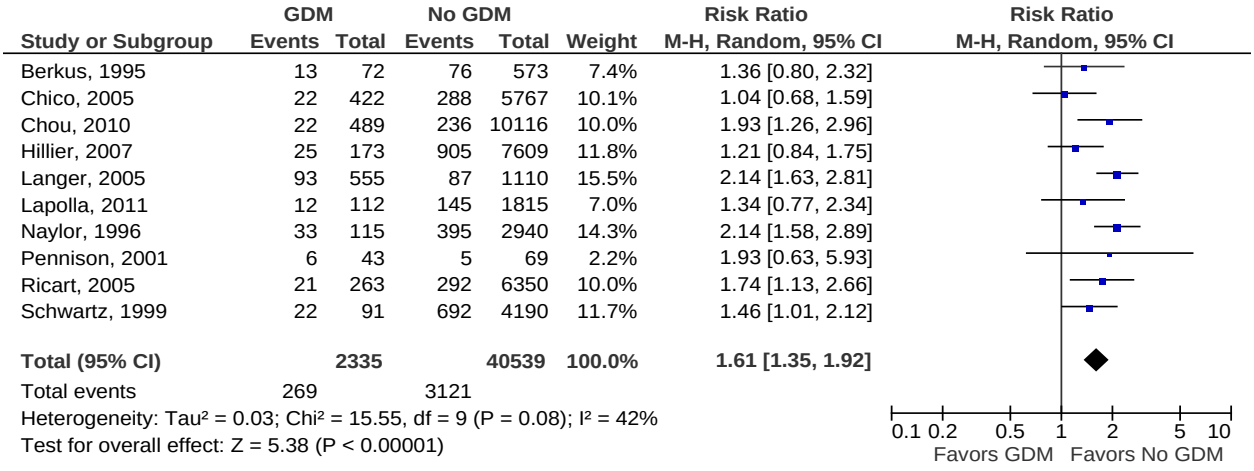
Short Term

A summary of the evidence for short and long term fetal, neonatal, and child outcomes is found in Table 14. The strength of evidence is presented in Table 15. The sections that follow describe the results by outcome.

Macrosomia (>4,000 g)

Twenty-one studies presented data for macrosomia (over 4,000 g) (Table 14).[79,80,84-86,88-90,92,93,102,106,132,133,135,136,145,146,150,152,160] There were significantly fewer cases of macrosomia in the patient groups with no GDM compared with CC GDM (10 studies, Figure 27).[86,89,92,93,102,132,136,146,150,160] CC 1 abnormal OGTT (7 studies, Figure 28),[80,86,106,132,135,136,145] NDDG GDM (1 study),[152] NDDG false positives (4 studies, Figure 29),[83,86,88,132] and WHO IGT (1 study).[133] Significantly fewer cases of macrosomia were observed among women with false-positive results compared with CC GDM (5 studies, Figure 30).[90,102,132,150,160] There was no significant difference in other comparisons involving other CC groups (Figure 31, Figure 32, Figure 33). One study compared WHO GDM with no GDM; no significant difference was observed between groups.[84] Two studies compared women who met IADPSG criteria for GDM with a no GDM group; no difference was observed between groups (Figure 34).[79,93] The strength of evidence for this outcome was low to insufficient due to risk of bias (all observational studies), inconsistency across studies, and/or imprecision in effect estimates (Table 15).

Figure 27. CC GDM versus no GDM: macrosomia (>4,000 g)

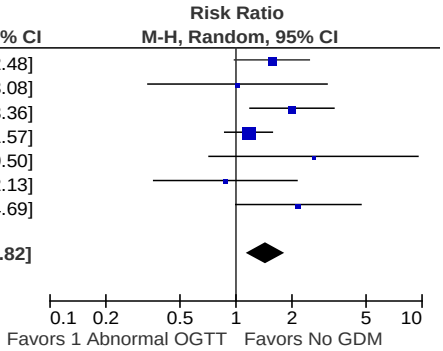


CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Figure 28. CC, 1 abnormal OGTT versus no GDM: macrosomia (>4,000 g)

Study or Subgroup	1 Abnormal OGTT Events	Total	No GDM Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Berkus, 1995	18	87	76	573	20.8%	1.56 [0.98, 2.48]
Chico, 2005	3	59	288	5767	4.4%	1.02 [0.34, 3.08]
Corrado, 2009	19	152	39	624	17.2%	2.00 [1.19, 3.36]
Hillier, 2007	40	288	905	7609	39.0%	1.17 [0.87, 1.57]
Lapolla, 2007	3	48	8	334	3.3%	2.61 [0.72, 9.50]
Rust, 1996	6	78	18	205	6.7%	0.88 [0.36, 2.13]
Vambergue, 2000	21	131	8	108	8.6%	2.16 [1.00, 4.69]
Total (95% CI)		**843**		**15220**	**100.0%**	**1.44 [1.13, 1.82]**
Total events	110		1342			

Heterogeneity: Tau² = 0.02; Chi² = 6.99, df = 6 (P = 0.32); I² = 14%
Test for overall effect: Z = 2.99 (P = 0.003)

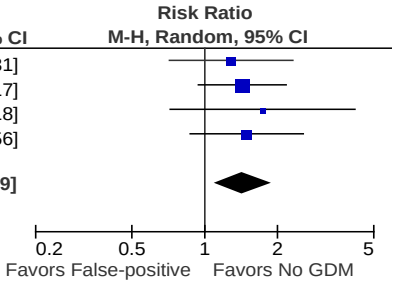


CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; OGTT = oral glucose tolerance test

Figure 29. NDDG false positive versus no GDM: macrosomia (>4,000 g)

Study or Subgroup	NDDG False-positive Events	Total	No GDM Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Chico, 2005	15	187	33	529	21.6%	1.29 [0.71, 2.31]
Hillier, 2007	27	326	83	1432	42.9%	1.43 [0.94, 2.17]
Retnakaran, 2008	18	128	6	74	9.7%	1.73 [0.72, 4.18]
Stamilio, 2004	14	164	95	1661	25.8%	1.49 [0.87, 2.56]
Total (95% CI)		**805**		**3696**	**100.0%**	**1.44 [1.10, 1.89]**
Total events	74		217			

Heterogeneity: Tau² = 0.00; Chi² = 0.33, df = 3 (P = 0.95); I² = 0%
Test for overall effect: Z = 2.61 (P = 0.009)

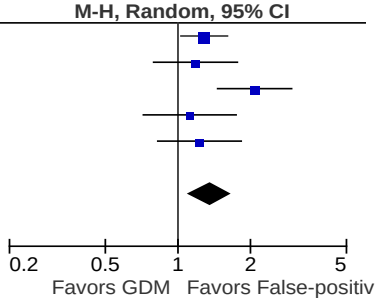


CI = confidence interval; GDM = gestational diabetes mellitus; NDDG = National Diabetes Data Group; M-H = Mantel-Haenszel

Figure 30. CC GDM versus false positive: macrosomia (>4,000 g)

Study or Subgroup	CC GDM Events	Total	False-positive Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Berggren, 2011	78	460	411	3117	30.1%	1.29 [1.03, 1.60]
Hillier, 2007	25	173	122	999	17.3%	1.18 [0.79, 1.76]
Naylor, 1996	33	115	80	580	20.0%	2.08 [1.46, 2.96]
Ricart, 2005	21	263	131	1838	15.2%	1.12 [0.72, 1.74]
Schwartz, 1999	22	91	119	605	17.4%	1.23 [0.83, 1.83]
Total (95% CI)		**1102**		**7139**	**100.0%**	**1.36 [1.10, 1.68]**
Total events	179		863			

Heterogeneity: Tau² = 0.03; Chi² = 7.33, df = 4 (P = 0.12); I² = 45%
Test for overall effect: Z = 2.82 (P = 0.005)

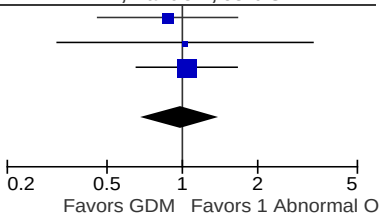


CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Figure 31. CC GDM versus 1 Abnormal OGTT: macrosomia (>4,000 g)

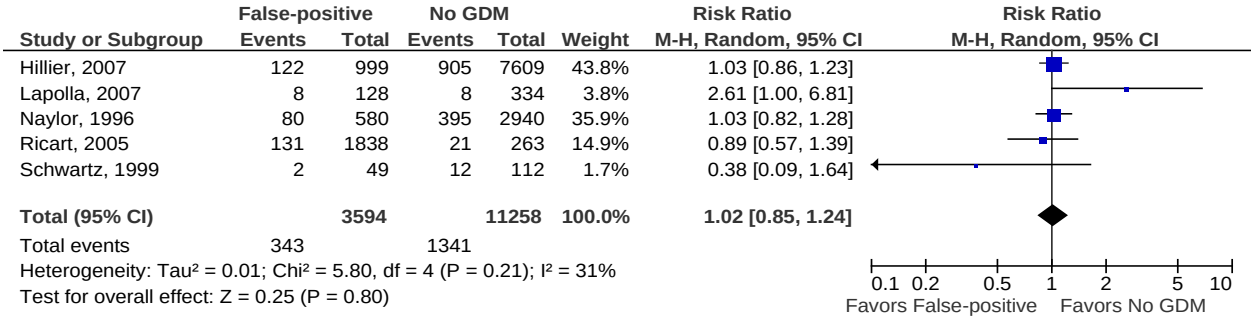
Study or Subgroup	GDM Events	Total	1 Abnormal OGTT Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Berkus, 1995	13	72	18	87	31.1%	0.87 [0.46, 1.66]
Chico, 2005	22	422	3	59	9.3%	1.03 [0.32, 3.32]
Hillier, 2007	25	173	40	288	59.7%	1.04 [0.66, 1.65]
Total (95% CI)		**667**		**434**	**100.0%**	**0.98 [0.69, 1.41]**
Total events	60		61			

Heterogeneity: Tau² = 0.00; Chi² = 0.20, df = 2 (P = 0.91); I² = 0%
Test for overall effect: Z = 0.09 (P = 0.93)



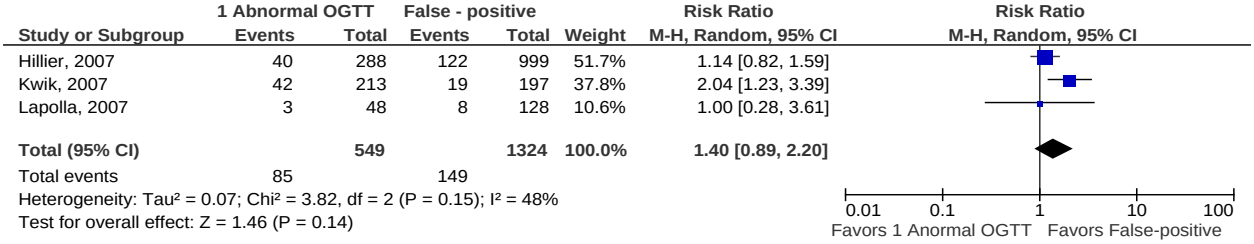
CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; OGTT = oral glucose tolerance test

Figure 32. CC false positives versus no GDM: macrosomia (>4,000 g)



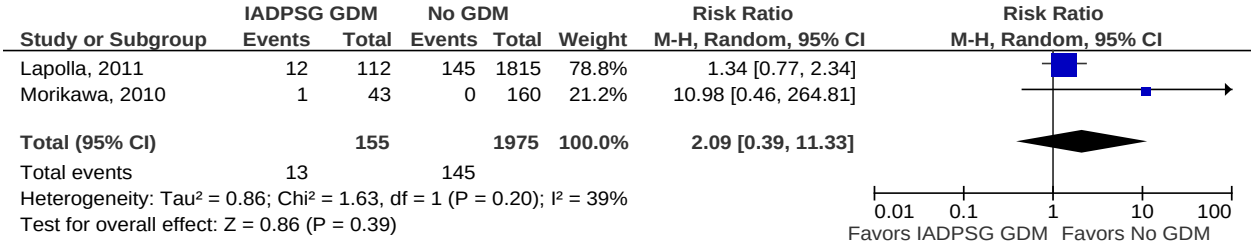
CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Figure 33. CC, 1 Abnormal OGTT versus False positives: macrosomia (>4,000 g)



CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Figure 34. IADPSG GDM versus No GDM: macrosomia (>4,000 g)

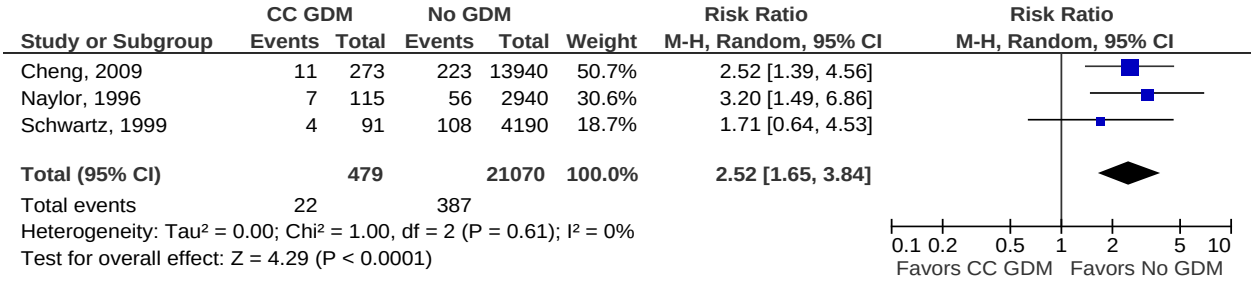


CI = confidence interval; GDM = gestational diabetes mellitus; IADPSG = International Association of the Diabetes in Pregnancy Study Groups; M-H = Mantel-Haenszel

Macrosomia (>4,500 g)

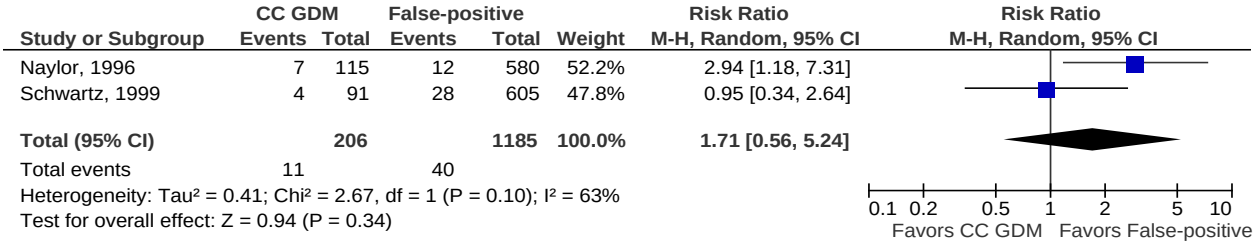
Four studies presented data on macrosomia (over 4,500 g) (Table 14).[81,150,152,160] Three studies showed a significant difference favoring the group with no GDM compared with CC GDM (Figure 35). The strength of evidence for this finding was low. No significant difference was found for CC GDM compared with false positives (2 studies; Figure 36) and CC false positives versus groups with no GDM (2 studies; Figure 37). One study compared NDDG GDM with a no GDM group, and found a significant difference in favor of the no GDM group.[152] The strength of evidence for these three findings was insufficient (Table 15).

Figure 35. CC GDM versus no GDM: macrosomia (>4,500 g)



CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Figure 36. CC GDM versus false positive: macrosomia (>4,500 g)

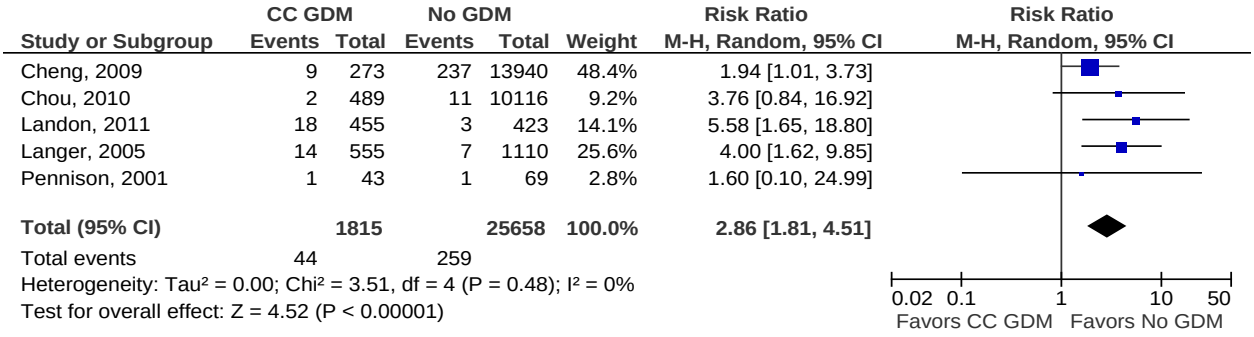


CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Shoulder Dystocia

Twelve studies presented data on shoulder dystocia (Table 14).[54,78,81,85,88-90,92,106,133,146,152] Five studies compared women who met CC criteria for GDM with no GDM and found a significant difference in favor of the no GDM group (Figure 37); the strength of evidence was rated low (Table 15).[81,89,92,146,163] One study compared CC GDM with a false-positive group, no significant difference was noted.[90] One study compared one abnormal OGTT by CC criteria with no GDM and no significant difference was found between groups.[106] One study compared women with 1 abnormal OGTT value by CC criteria with a false-positive group with a significant difference noted in favor of the false-positive group.[85] One study compared unrecognized GDM by NDDG criteria with a no GDM group;[152] another study compared a false-positive group with no GDM.[88] Both studies noted a significant difference in favor of the groups with no GDM. A single study compared IGT by WHO criteria and no GDM; a significant difference was found in favor of group with no GDM.[133] One study compared varying degrees of glucose intolerance by IADPSG criteria and no GDM;[78] significant differences were observed when no GDM was compared with IFG and IGT and fasting glucose combined. No GDM was favored in both cases. The remaining groups demonstrated no significant differences (Table 14). The strength of evidence for all comparisons based on single studies was rated insufficient (Table 15).

Figure 37. CC GDM versus no GDM: shoulder dystocia



CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Clavicular Fracture

No studies provided comparable data on clavicular fracture. However, this outcome was often a composite outcome within birth injury or fetal birth trauma.

Brachial Plexus Injury

No studies provided comparable data on brachial plexus injury, also often a composite of birth injury or fetal birth trauma.

Fetal Birth Trauma or Birth Injury

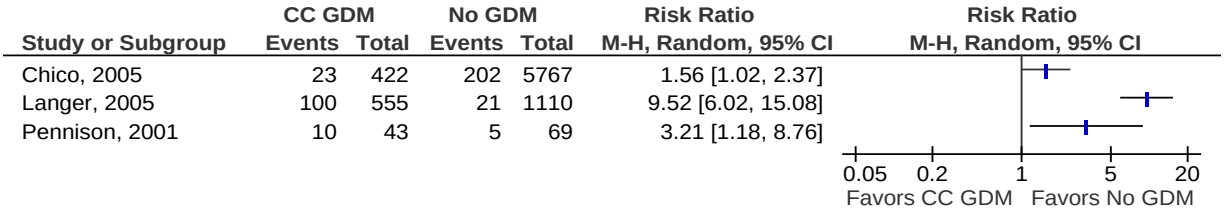
Four studies presented data for fetal birth trauma or traumatic delivery (Table 14).[81,149,152,155] Birth trauma was undefined in two studies,[149,155] one comparing WHO IGT with no GDM. Another defined birth trauma as a composite of brachial plexus injury, facial nerve palsy, clavicular fracture, skull fracture, and head laceration; this study compared CC GDM and no GDM.[81] No significant difference was observed in any comparison. Brachial plexus injury, cranial nerve palsy, and clavicular facture were also components of birth trauma in one study.[152] This study compared women with unrecognized NDDG GDM and no GDM and showed a significant difference in favor of the no GDM group. Strength of evidence for all comparisons was insufficient.

Hypoglycemia

Twelve studies presented data on neonatal hypoglycemia (Table 14).[67,80,86,89,103,106,133,135,146,149,152,155] Two studies did not define hypoglycemia,[67,125] while all other studies defined hypoglycemia with varying glucose threshold criteria or by necessity of intravenous glucose. Three studies compared women meeting CC criteria for GDM with groups without GDM. Results were not pooled due to substantial heterogeneity across studies (I^2=94%) (Figure 38); however, all three studies individually showed fewer cases of hypoglycemia among the patient groups with no GDM.[86,89,146] The difference in results may be explained in part by the methods of assessing for neonatal hypoglycemia (e.g., biochemical vs. clinical). Posthoc analysis showed that the magnitude of association between glucose intolerance and neonatal hypoglycemia was affected by the definition used (i.e., clinical or biochemical). Many of the observational studies included did not routinely apply the same biochemical screening procedure to the non-GDM groups and glucose intolerant women. No significant difference was found for remaining comparisons. One study compared women meeting CC criteria for GDM with women demonstrating one abnormal OGTT value,[86] and four studies compared women meeting CC criteria on one abnormal OGTT value with no GDM (Figure 39).[80,86,106,135] One study compared

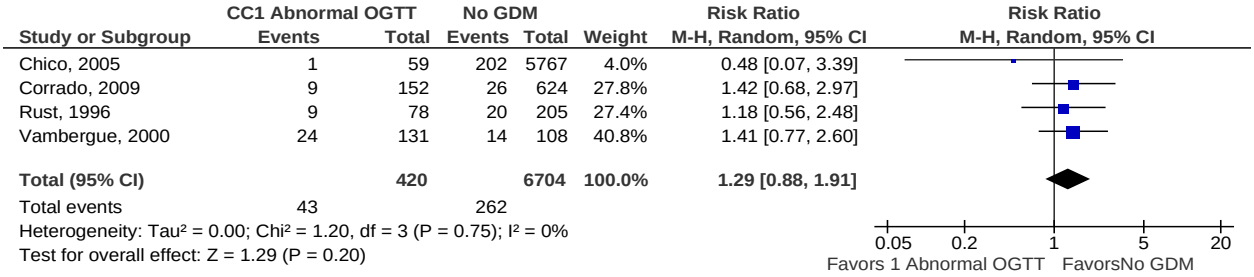
women who met NDDG criteria for GDM with no GDM,[152] one study compared NDDG false positive with no GDM,[67] and another study compared NGGD 1 abnormal OGTT versus no GDM.[103] Three studies compared women meeting WHO criteria for IGT with no GDM (Figure 40).[133,149] Strength of evidence for all comparisons was insufficient.

Figure 38. CC GDM versus no GDM: hypoglycemia



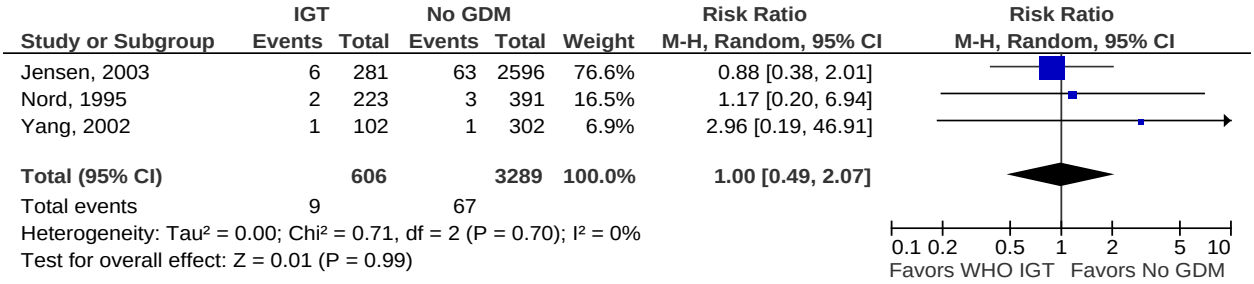
CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Figure 39. CC, 1 abnormal OGTT versus no GDM: hypoglycemia



CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; OGTT = oral glucose tolerance tests

Figure 40. WHO impaired glucose tolerance versus no GDM: hypoglycemia



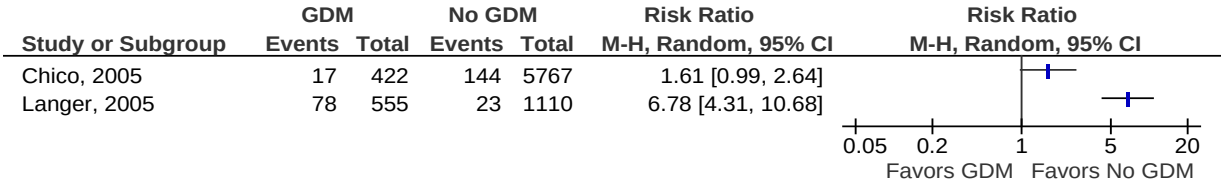
CI = confidence interval; GDM = gestational diabetes mellitus; IGT = impaired glucose tolerance; M-H = Mantel-Haenszel; WHO = World Health Organization

Hyperbilirubinemia

Eight studies presented data for hyperbilirubinemia or neonatal jaundice (Table 14).[67,78,86,87,106,133,146,149] Plasma bilirubin values for the diagnosis of hyperbilirubinemia varied amongst studies. Of the seven studies, four studies compared differing CC criterion, including CC GDM with no GDM (Figure 41),[86,146] CC GDM and one abnormal OGTT,[86] CC 1 abnormal OGTT and no GDM,[106] and CC false positive and no GDM.[87] Results for CC GDM versus no GDM were not pooled due to substantial statistical heterogeneity (I^2=94%). Possible sources of heterogeneity include differences in assessing outcomes across studies and uncontrolled differences between comparison groups. CC false positive versus no GDM showed a significant difference with fewer cases in the group with no GDM. The other comparison involving CC criteria (CC GDM vs. 1 abnormal OGTT) showed no significant difference between groups. One

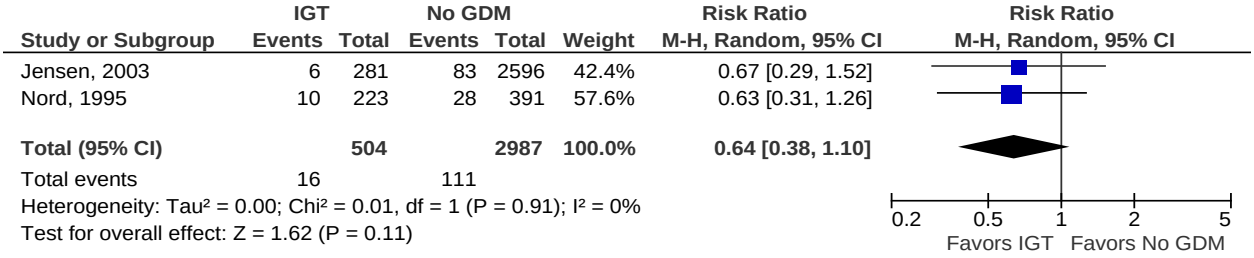
study compared women with a false-positive result by NDDG criteria with no GDM; no significant difference was found.[67] Two studies compared women meeting WHO criteria for IGT with no GDM; no significant difference was found (Figure 42).[133,149] One study compared various IADPSG thresholds for glucose intolerance.[78] A significant difference was present in comparisons of IADPSG isolated (1 value above threshold) IGT and double-isolated (two values above threshold) IGT with no GDM, both favoring the no GDM group. No further differences were observed for any other IADPSG comparisons.

Figure 41. CC GDM versus no GDM: hyperbilirubinemia



CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Figure 42. WHO impaired glucose tolerance versus no GDM: hyperbilirubinemia



CI = confidence interval; GDM = gestational diabetes mellitus; IGT = impaired glucose tolerance; M-H = Mantel-Haenszel; WHO = World Health Organization

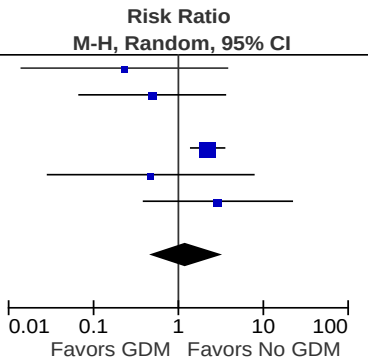
Morbidity/Mortality

Sixteen studies presented data for neonatal mortality or morbidity (Table 14).[67,85-88,92,93,102,103,106,135,146,149,150,154,155] No studies demonstrated a significant difference between groups which may be due to small numbers of events within some comparisons. Six studies compared women meeting CC criteria for GDM with no GDM (Figure 43),[86,92,93,102,146,150] two studies compared CC GDM with false positives (Figure 44),[102,150] and one study compared women with CC GDM and those with one abnormal OGTT.[86] Three studies compared one abnormal OGTT to no GDM (Figure 45),[86,106,135] three studies compared women with false-positive results by CC criteria with no GDM (Figure 46),[87,102,150] and one study compared CC false positive with one abnormal OGTT value.[85] Two studies compared women with false-positive results by NDDG criteria with no GDM (Figure 47),[67,88] one study compared NDDG 1 abnormal OGTT versus no GDM,[103] three studies employed WHO criteria for IGT compared with no GDM (Figure 48),[149,154,155] and another study followed IADPSG criteria for GDM diagnosis compared with no GDM.[93]

Figure 43. CC GDM versus no GDM: morbidity/mortality

Study or Subgroup	GDM Events	GDM Total	No GDM Events	No GDM Total	Weight	Risk Ratio M-H, Random, 95% CI
Chico, 2005	0	422	29	5767	10.1%	0.23 [0.01, 3.78]
Chou, 2010	1	489	42	10116	16.8%	0.49 [0.07, 3.57]
Langer, 2005	0	555	0	1110		Not estimable
Lapolla, 2011	18	112	132	1815	46.6%	2.21 [1.40, 3.48]
Ricart, 2005	0	263	25	6350	10.1%	0.47 [0.03, 7.73]
Schwartz, 1999	1	154	16	7207	16.4%	2.92 [0.39, 21.92]
Total (95% CI)		**1995**		**32365**	**100.0%**	**1.23 [0.46, 3.30]**
Total events	20		244			

Heterogeneity: Tau² = 0.49; Chi² = 6.62, df = 4 (P = 0.16); I² = 40%
Test for overall effect: Z = 0.40 (P = 0.69)

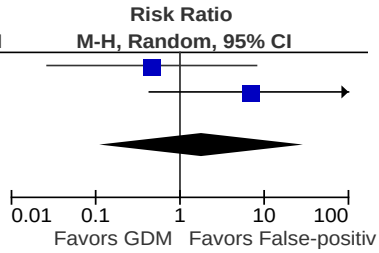


CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Figure 44. CC GDM versus false positive: morbidity/mortality

Study or Subgroup	GDM Events	GDM Total	False-positive Events	False-positive Total	Weight	Risk Ratio M-H, Random, 95% CI
Ricart, 2005	0	263	7	1838	49.1%	0.46 [0.03, 8.11]
Schwartz, 1999	1	154	1	1066	50.9%	6.92 [0.44, 110.10]
Total (95% CI)		**417**		**2904**	**100.0%**	**1.83 [0.11, 29.41]**
Total events	1		8			

Heterogeneity: Tau² = 1.95; Chi² = 1.95, df = 1 (P = 0.16); I² = 49%
Test for overall effect: Z = 0.43 (P = 0.67)

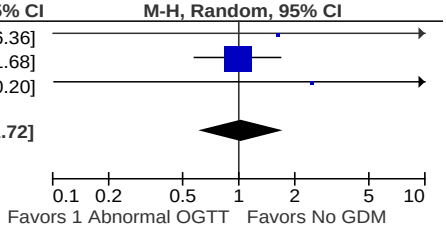


CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Figure 45. CC, 1 abnormal OGTT versus no GDM: morbidity/mortality

Study or Subgroup	1 Abnormal OGTT Events	1 Abnormal OGTT Total	No GDM Events	No GDM Total	Weight	Risk Ratio M-H, Random, 95% CI
Chico, 2005	0	59	29	5767	3.4%	1.63 [0.10, 26.36]
Rust, 1996	15	78	40	205	93.9%	0.99 [0.58, 1.68]
Vambergue, 2000	1	131	0	108	2.6%	2.48 [0.10, 60.20]
Total (95% CI)		**268**		**6080**	**100.0%**	**1.03 [0.61, 1.72]**
Total events	16		69			

Heterogeneity: Tau² = 0.00; Chi² = 0.42, df = 2 (P = 0.81); I² = 0%
Test for overall effect: Z = 0.10 (P = 0.92)

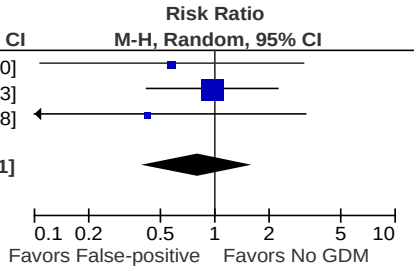


CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; OGTT = oral glucose tolerance test

Figure 46. CC false positive versus no GDM: morbidity/mortality

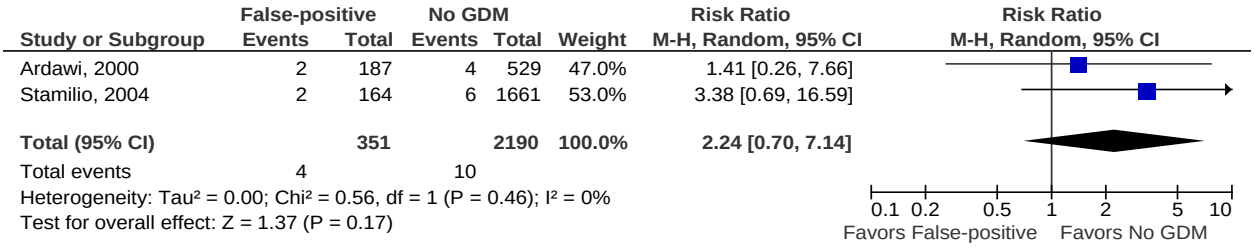
Study or Subgroup	False-positive Events	False-positive Total	No GDM Events	No GDM Total	Weight	Risk Ratio M-H, Random, 95% CI
Bo, 2004	4	315	2	91	17.4%	0.58 [0.11, 3.10]
Ricart, 2005	7	1838	25	6350	70.5%	0.97 [0.42, 2.23]
Schwartz, 1999	1	1066	16	7207	12.1%	0.42 [0.06, 3.18]
Total (95% CI)		**3219**		**13648**	**100.0%**	**0.80 [0.40, 1.61]**
Total events	12		43			

Heterogeneity: Tau² = 0.00; Chi² = 0.73, df = 2 (P = 0.69); I² = 0%
Test for overall effect: Z = 0.62 (P = 0.53)



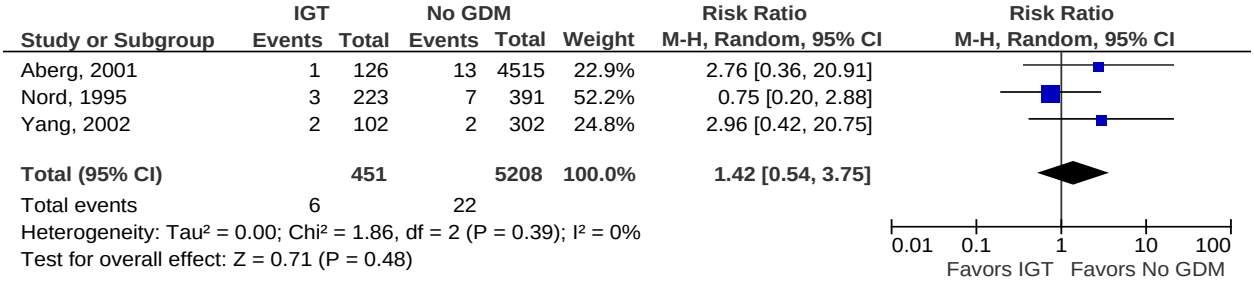
CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel

Figure 47. NDDG false positive versus no GDM: morbidity/mortality



CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; NDDG = National Diabetes Data Group

Figure 48. WHO IGT versus no GDM: morbidity/mortality



CI = confidence interval; GDM = gestational diabetes mellitus; IGT = impaired glucose tolerance; M-H = Mantel-Haenszel; WHO = World Health Organization

Long Term

One study presented data on long term health outcomes for infants and children (i.e., prevalence of childhood obesity).[132]

Prevalence of Childhood Obesity

Significant differences were found between women meeting thresholds for CC GDM in comparison with those without GDM, favoring the no GDM group.[132] The CC false-positive group was favored compared with women meeting CC GDM criteria (Table 14). These findings should be interpreted cautiously because this study did not adjust for maternal BMI, one of the most important predictors of childhood obesity. No significant differences were found for the remaining comparisons (Table 14).

Table 14. Evidence summary table: fetal/neonatal outcomes

Outcome	Comparison	Studies	Participants	Effect Estimate*	I^2	Favors‡
Macrosomia >4,000 g	CC GDM vs. no GDM	10	42,874	1.61 [1.35, 1.92]	42%	No GDM
	CC GDM vs. false positive	5	8,241	1.36 [1.10, 1.68]	45%	False positive
	CC GDM vs. 1 abnormal OGTT	3	1,101	0.98 [0.69, 1.41]	0%	-
	CC 1 abnormal OGTT vs. no GDM	7	16,063	1.44 [1.13, 1.82]	14%	No GDM
	CC false positive vs. no GDM	5	14,852	1.02 [0.85, 1.24]	31%	-
	CC 1 abnormal OGTT vs. false positive	3	1,873	1.40 [0.89, 2.20]	48%	-
	NDDG GDM (unrecognized) vs. no GDM	1	80	5.60 [2.04, 15.35]	NA	No GDM
	NDDG false positive vs. no GDM	4	4,501	1.44 [1.10, 1.89]	0%	No GDM
	WHO GDM vs. no GDM	1	542	3.33 [0.49, 22.70]	NA	-
	WHO IGT vs. no GDM	1	2,885	1.26 [1.06, 1.50]	NA	No GDM
	IADPSG GDM vs. no GDM	2	2,130	2.09 [0.39, 11.33]	39%	-
Macrosomia >4,500 g	CC GDM vs. no GDM	3	21,549	2.52 [1.65, 3.84]	0%	No GDM
	CC GDM vs. false positive	2	1,391	1.71 [0.56, 5.24]	63%	-
	CC false positive vs. no GDM	2	8,315	1.48 [0.91, 2.39]	44%	-
	NDDG GDM (unrecognized) vs. no GDM	1	80	26.76 [1.45, 493.62]	NA	No GDM

Table 14. Evidence summary table: fetal/neonatal outcomes (continued)

Outcome	Comparison	Studies	Participants	Effect Estimate*	I^2	Favors‡
Shoulder dystocia	CC GDM vs. no GDM	5	27,473	2.86 [1.81, 4.51]	0%	No GDM
	CC GDM vs. false positive	1	3,577	1.49 [0.97, 2.30]	NA	-
	CC 1 abnormal OGTT vs. no GDM	1	239	0.20 [0.02, 1.82]	NA	-
	CC 1 abnormal OGTT vs. false positive	1	410	5.09 [1.14, 22.66]	NA	False positive
	NDDG GDM (unrecognized) vs. no GDM	1	80	6.00 [1.09, 32.95]	NA	No GDM
	NDDG false positive vs. no GDM	1	1,825	2.79 [1.30, 6.01]	NA	No GDM
	WHO IGT vs. no GDM	1	2,885	2.18 [1.02, 4.67]	NA	No GDM
	IADPSG IGT vs. no GDM	1	7,411	1.21 [0.76, 1.92]	NA	-
	IADPSG IFG vs. no GDM	1	7,906	1.48 [1.10, 1.98]	NA	No GDM
	IADPSG IGT-2 vs. no GDM	1	7,103	1.58 [0.67, 3.72]	NA	-
	IADPSG IGT IFG vs. no GDM	1	7,351	1.82 [1.21, 2.75]	NA	No GDM
	IADPSG IGT vs. IFG	1	1,277	0.82 [0.48, 1.38]	NA	-
	IADPSG IGT vs. IGT-2	1	474	0.76 [0.29, 2.00]	NA	-
	IADPSG IGT vs. IGT IFG	1	722	0.66 [0.36, 1.21]	NA	-
	IADPSG IFG vs. IGT-2	1	969	0.94 [0.38, 2.28]	NA	-
	IADPSG IFG vs. IGT IFG	1	1,217	0.81 [0.50, 1.31]	NA	-
	IADPSG IGT-2 vs. IGT IFG	1	414	0.87 [0.34, 2.21]	NA	-

Table 14. Evidence summary table: fetal/neonatal outcomes (continued)

Outcome	Comparison	Studies	Participants	Effect Estimate*	I^2	Favors‡
Neonatal hypoglycemia	CC GDM vs. no GDM	3	7,966	Results not pooled due to substantial heterogeneity.	94%	-
	CC GDM vs. 1 abnormal OGTT	1	481	3.22 [0.44, 23.37]	NA	-
	CC 1 abnormal OGTT vs. no GDM	4	7,124	1.29 [0.88, 1.91]	0%	-
	NDDG GDM vs. no GDM	1	80	Not Estimable†	NA	-
	NDDG false positive vs. no GDM	1	716	2.83 [0.58, 13.89]	NA	-
	NDDG, 1 abnormal OGTT vs. no GDM	1	699	9.60 [0.86, 106.73]	NA	-
	WHO IGT vs. no GDM	3	3,895	1.00 [0.49, 2.07]	0%	-
Hyperbilirubinemia	CC GDM vs. no GDM	2	7,854	Results not pooled due to substantial heterogeneity.	94%	-
	CC GDM vs. 1 abnormal OGTT	1	481	2.38 [0.32, 17.53]	NA	-
	CC false positive vs. no GDM	1	406	3.03 [1.12, 8.23]	NA	No GDM
	CC 1 abnormal OGTT vs. no GDM	1	239	4.19 [0.20, 88.20]	NA	-
	NDDG False positive vs. no GDM	1	716	1.07 [0.68, 1.70]	NA	-
	WHO IGT vs. no GDM	2	3,491	0.64 [0.38, 1.10]	0%	-
	IADPSG IGT vs. no GDM	1	7,411	1.32 [1.06, 1.64]	NA	No GDM
	IADPSG IFG vs. no GDM	1	7,906	1.03 [0.87, 1.23]	NA	-
	IADPSG IGT-2 vs. no GDM	1	7,103	1.55 [1.03, 2.35]	NA	No GDM

Table 14. Evidence summary table: fetal/neonatal outcomes (continued)

Outcome	Comparison	Studies	Participants	Effect Estimate*	I^2	Favors‡
Hyperbilirubinemia (continued)	IADPSG IGT IFG vs. no GDM	1	7,351	0.97 [0.74, 1.29]	NA	-
	IADPSG IGT vs. IFG	1	1,277	1.27 [0.98, 1.66]	NA	-
	IADPSG IGT vs. IGT-2	1	474	0.85 [0.54, 1.34]	NA	-
	IADPSG IGT vs. IGT IFG	1	722	1.35 [0.96, 1.91]	NA	-
	IADPSG IFG vs. IGT-2	1	969	0.67 [0.43, 1.03]	NA	-
	IADPSG IFG vs. IGT IFG	1	1,217	1.06 [0.78, 1.46]	NA	-
	IADPSG IGT-2 vs. IGT IFG	1	414	1.60 [0.98, 2.61]	NA	-
Fetal birth trauma/injury	CC GDM vs. no GDM	1	14,213	1.19 [0.68, 2.08]	NA	-
	NDDG GDM vs. no GDM	1	80	34.41 [1.95, 608.47]	NA	No GDM
	WHO IGT vs. no GDM	2	1,018	0.29 [0.04, 2.41]	NA	-
Fetal morbidity/mortality	CC GDM vs. no GDM	6	34,360	1.23 [0.46, 3.30]	40%	-
	CC GDM vs. false positive	2	3,321	1.83 [0.11, 29.41]	49%	-
	CC GDM vs. 1 abnormal OGTT	1	481	Not estimable†	NA	-
	CC 1 abnormal OGTT vs. no GDM	3	6,348	1.03 [0.61, 1.72]	0%	-
	CC false positive vs. no GDM	3	16,867	0.80 [0.40, 1.61]	0%	-
	CC false positive vs. 1 abnormal OGTT	1	410	Not estimable†	NA	-
	NDDG false positive vs. no GDM	2	2,541	2.24 [0.70, 7.14]	0%	-

Table 14. Evidence summary table: fetal/neonatal outcomes (continued)

Outcome	Comparison	Studies	Participants	Effect Estimate*	I^2	Favors‡
Fetal morbidity/mortality (continued)	NDDG 1 abnormal OGTT vs. no GDM	1	699	0.94 [0.04, 19.69]	NA	-
	WHO IGT vs. no GDM	3	5,659	1.42 [0.54,3.75]	0%	-
	IADPSG GDM vs. no GDM	1	1927	2.21 [1.40, 3.48]	NA	-
Prevalence of childhood obesity	CC GDM vs. no GDM	1	7,782	1.48 [1.20, 1.82]	NA	No GDM
	CC GDM vs. false positive	1	1,172	1.49 [1.18,1.88]	NA	False positive
	CC GDM vs. 1 abnormal OGTT	1	461	1.30 [0.98, 1.72]	NA	-
	CC false positive vs. no GDM	1	8,608	0.99 [0.88, 1.12]	NA	-
	CC false positive vs. 1 abnormal OGTT	1	1,287	0.81 [0.56, 1.18]	NA	-
	CC 1 abnormal OGTT vs. no GDM	1	7,897	1.14 [0.94, 1.38]	NA	-

CC = Carpenter-Coustan; GDM = gestational diabetes mellitus; IFG = impaired fasting glycemia; IGT = impaired glucose tolerance; IGT-2 = double impaired glucose tolerance; IADPSG = International Association of Diabetes and Pregnancy Study Groups; NDDG = National Diabetes Data Group; NA = not applicable; OGTT = oral glucose tolerance test; WHO = World Health Organization.

*****Effect estimates are risk ratios with 95% confidence intervals.

†Not estimable due to zero events in both groups.

‡Where the result was statistically significant, we have listed the group that had the better outcome (e.g., lower incidence of macrosomia).

Table 15. Strength of evidence summary table: fetal/neonatal outcomes

Outcome	Comparison	Studies	Risk of Bias	Consistency	Directness	Precision	SOE
Macrosomia >4,000 g	CC GDM vs. no GDM	10	High	Consistent	Direct	Precise	Low
	CC GDM vs. false positive	5	High	Consistent	Direct	Precise	Low
	CC GDM vs. 1 abnormal OGTT	3	High	Consistent	Direct	Precise	Low
	CC 1 abnormal OGTT vs. no GDM	7	High	Consistent	Direct	Precise	Low
	CC false positive vs. no GDM	5	High	Consistent	Direct	Precise	Low
	CC 1 abnormal OGTT vs. false positive	3	High	Inconsistent	Direct	Precise	Insufficient
	NDDG GDM (unrecognized) vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	NDDG false positive vs. no GDM	4	High	Consistent	Direct	Precise	Low
	WHO GDM vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	WHO IGT vs. no GDM	1	high	Unknown	Direct	Precise	Insufficient
	IADPSG GDM vs. no GDM	2	High	Consistent	Direct	Imprecise	Insufficient
Macrosomia >4,500 g	CC GDM vs. no GDM	3	High	Consistent	Direct	Precise	Low
	CC GDM vs. false positive	2	High	Inconsistent	Direct	Imprecise	Insufficient
	CC false positive vs. no GDM	2	High	Consistent	Direct	Imprecise	Insufficient
	NDDG GDM (unrecognized) vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
Shoulder dystocia	CC GDM vs. no GDM	5	High	Consistent	Direct	Precise	Low
	CC GDM vs. false positive	1	High	Unknown	Direct	Imprecise	Insufficient
	CC 1 abnormal OGTT vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	CC 1 abnormal OGTT vs. false positive	1	High	Unknown	Direct	Imprecise	Insufficient
	NDDG GDM (unrecognized) vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	NDDG false positive vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	WHO IGT vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IGT vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IFG vs. no GDM	1	High	Unknown	Direct	Precise	Insufficient
	IADPSG IGT-2 vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IGT IFG vs. no GDM	1	High	Unknown	Direct	Precise	Insufficient

Table 15. Strength of evidence summary table: fetal/neonatal outcomes (continued)

Outcome	Comparison	Studies	Risk of Bias	Consistency	Directness	Precision	SOE
Shoulder dystocia (continued)	IADPSG IGT vs. IFG	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IGT vs. IGT-2	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IGT vs. IGT IFG	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IFG vs. IGT-2	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IFG vs. IGT IFG	1	High	Unknown	Direct	Imprecise	Insufficient
	IADPSG IGT-2 vs. IGT IFG	1	High	Unknown	Direct	Imprecise	Insufficient
Neonatal hypoglycemia	CC GDM vs. no GDM	3	High	Inconsistent	Direct	Imprecise	Insufficient
	CC GDM vs. 1 abnormal OGTT	1	High	Unknown	Direct	Imprecise	Insufficient
	CC 1 abnormal OGTT vs. no GDM	4	High	Consistent	Direct	Imprecise	Insufficient
	NDDG GDM vs. no GDM	1	High	Unknown	Direct	NA	Insufficient
	NDDG false positive vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	NDDG 1 abnormal OGTT vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	WHO IGT vs. no GDM	3	High	Consistent	Direct	Imprecise	Insufficient
Fetal birth trauma/injury	CC GDM vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	NDDG GDM vs. no GDM	1	High	Unknown	Direct	Imprecise	Insufficient
	WHO IGT vs. no GDM	2	High	Consistent	Direct	Imprecise	Insufficient

CC = Carpenter-Coustan; GDM = gestational diabetes mellitus; IGT = impaired glucose tolerance; IGT-2 = double impaired glucose tolerance; IADPSG = International Association of Diabetes and Pregnancy Study Groups; IFG = impaired fasting glucose; NA = not applicable; NDDG = National Diabetes Data Group; OGTT = oral glucose tolerance test; SOE = strength of evidence; WHO = World Health Organization.

Key Question 4. Does treatment modify the health outcomes of mothers who meet various criteria for GDM and offspring?

Description of Included Studies

Eleven studies met the inclusion criteria for Key Question 4.[50,54,92,95-98,146,148,152,160] The studies are described in Appendix D. All studies compared diet modification, glucose monitoring and insulin as needed with standard care. Five of the studies were RCTs,[50,54,96-98] while six were retrospective cohort studies.[92,95,146,148,152,160] The studies were published between 1996 and 2010 (median year 2005). Two studies had two associated publications reporting initial and longer term outcomes.[50,54] Five studies were from the United States,[54,95,98,146,152] two from Italy,[97,148] two from Canada,[96,160] and one each from Taiwan[92] and Australia.[50] The screening test used in most studies was OGCT with a 100 g OGTT assessed using CC criteria, except for the studies from Canada and Australia that used a OGCT with a 75 g OGTT. Diagnostic testing in all studies occurred at or after 24 weeks' gestation. Among these studies a variety of glucose inclusion criteria were used, varying from 50 g screen positive with nondiagnostic OGTTs to women who met National Diabetes Data Group criteria for a diagnosis of GDM. The two largest RCTs[50,163] by Crowther et al. and Landon et al. used different glucose thresholds for entry in their trials: WHO and CC criteria with a fasting glucose <95 mg/dL (5.3 mmol/L), respectively; however, the mean glucose levels of women at study entry were remarkably similar between these two studies.

Methodological Quality of Included Studies

The methodological quality of the included studies is described in Appendix C3. The risk of bias for the RCTs was low for one trial,[50] unclear for three trials,[54,97,98] and high for one trial.[96] The trials that were unclear most commonly did not report detailed methods for sequence generation and allocation concealment. The trial assessed as high risk of bias was due to lack of blinding for outcome assessment and incomplete outcome data.

The six cohort studies were all considered high quality, with overall quality scores of 7 to 9 on a 9-point scale. Three studies received full scores of 9.[54,152,160] One study received a score of 8/9 because the study population was a selected (non-representative) group (i.e., participants at a diabetic center).[148] Two studies received a score of 7/9. One study obtained this score due to the study population considered only "somewhat" representative (all women were cared for under a single health plan); as well as a lack of control for potential confounders including age, race, BMI, previous GDM, or family history of DM.[95] The absence of control for any potential confounders was also the reason for the lower score in the second study. [92]

Key Points

- A variety of glucose threshold criteria were used for inclusion across studies. For outcomes where results were inconsistent between studies, different study glucose threshold entry criteria did not explain the variation.
- Results for some outcomes were driven by the two largest RCTs, the Maternal Fetal Medicine Unit (MFMU)[54] and the Australian Carbohydrate Intolerance in Pregnancy Study (ACHOIS),[50] which had unclear and low risk of bias, respectively.

Maternal Outcomes

- There was moderate evidence from 3 RCTs showing a significant difference for preeclampsia with fewer cases in the treated group.
- There was inconsistency across studies in terms of differences in maternal weight gain (4 RCTs and 2 cohort studies). The strength of evidence was considered insufficient due to inconsistency across studies and imprecision in effect estimates.
- No differences between groups were found for cesarean section (5 RCTs, 6 cohorts) or unplanned cesarean section (1 RCT, 1 cohort).
- There was inconsistency across studies in terms of induction of labor with no difference found for the 2 RCTs overall and a significant difference favoring the treatment group among the one cohort study included.
- Only one RCT reported on BMI at delivery and showed a significant difference with lower BMI in the treated group.
- Only one cohort study reported maternal birth trauma (i.e., postpartum hemorrhage) and showed no difference between groups.
- There was no evidence for long-term maternal outcomes such as type 2 diabetes mellitus, obesity, and hypertension.

Short-Term Outcomes in the Offspring

- There was insufficient evidence for birth injury. There was inconsistency across studies with the 2 RCTs showing no difference and the one cohort study showing a difference in favor of the treated group. The low number of events and participants across all studies resulted in imprecise estimates.
- The incidence of shoulder dystocia was significantly lower in the treated groups, and this finding was consistent for the 3 RCTs and 4 cohort studies. Overall, the evidence for shoulder dystocia was considered moderate showing a difference in favor of the treated group.
- For other injury outcomes, including brachial plexus injury (1 RCT, 1 cohort), and clavicular fractures (1 RCT, 1 cohort), the results were inconsistent across designs with the RCTs showing no differences between groups and the cohort study showing a significant difference in favor of the treated group.
- There was low evidence of no difference between groups for neonatal hypoglycemia based on four RCTs and 2 cohort studies.
- For outcomes related to birthweight (including macrosomia >4,000 g, macrosomia >4,500 g, actual birthweight, and large for gestational age), differences were often observed favoring the treated groups. The strength of evidence was moderate for macrosomia >4,000 g suggesting a benefit of treatment.
- There was no difference in hyperbilirubinemia for the 3 RCTs, while the one cohort study showed a significant difference in favor of the treated group.
- There were no differences observed across studies for perinatal death (3 RCTs, 3 cohorts). Two RCTs showed no difference between groups for respiratory distress syndrome, while one cohort study found a significant difference favoring the treated group for “respiratory complications.” Several studies assessed APGAR scores, and while differences were found in both the RCT and cohort study for APGAR at 1 minute, no differences were found among the 2 RCTs and 1 cohort study at 5 minutes.

Long-Term Outcomes in the Offspring

- One RCT followed patients for 7 to 11 years and found no differences for impaired glucose tolerance or type 2 diabetes mellitus, although the strength of evidence was considered insufficient.
- No differences were observed in single studies that assessed BMI >95 (7-11 year followup) and BMI >85 percentile (5-7 year followup). Overall, pooled results showed no difference in BMI and the strength of evidence was considered low.

Detailed Synthesis

Detailed results are described by outcome in the sections that follow. We first describe the maternal outcomes, followed by fetal/neonatal/child outcomes. We present meta-graphs when two or more studies were pooled. These are displayed after the description of results for each outcome. A detailed table of results is presented at the end of each of the maternal and fetal/neonatal/child sections (Table 16 and Table 17, respectively). The strength of evidence for key outcomes is presented in Table 18.

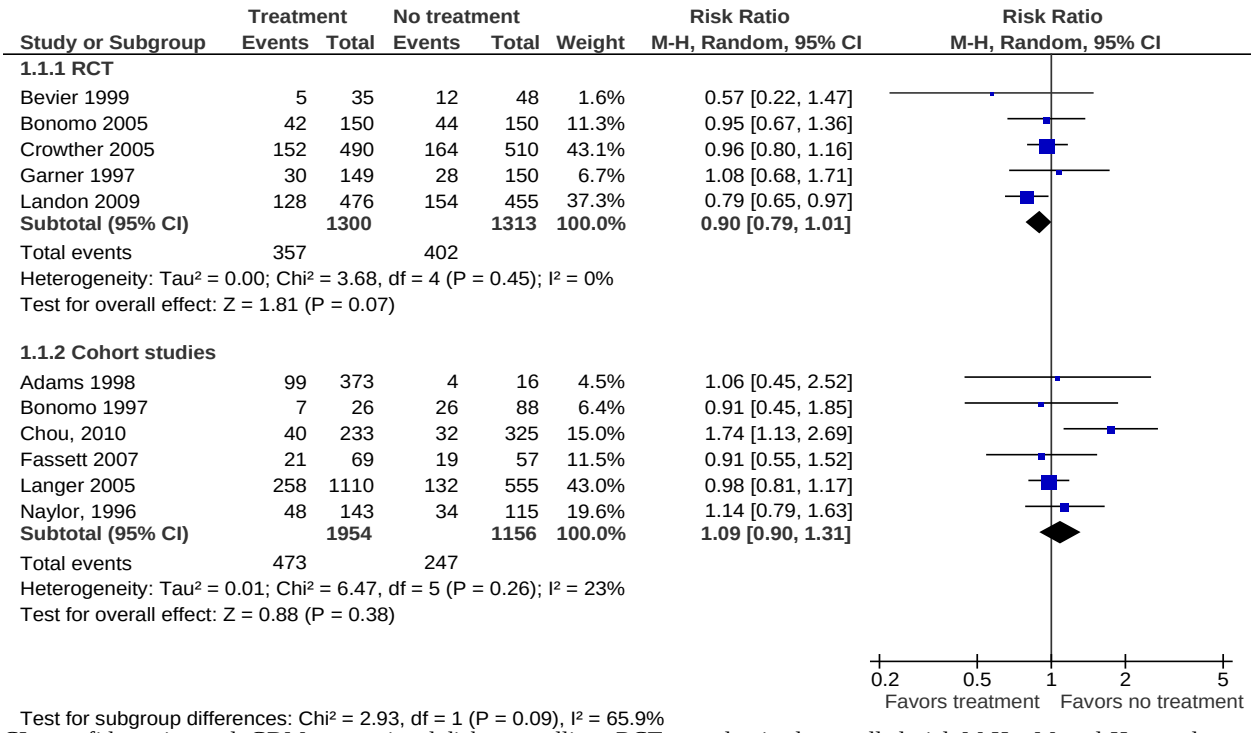
Maternal Outcomes

Short Term

Cesarean Delivery

All studies provided data on cesarean delivery (Table 16).[50,54,92,95-98,146,148,152,160] There was no significant difference in the pooled estimates for the RCTs (risk ratio [RR] 0.90, 95% CI 0.79 to 1.01, n = 2,613) or for the cohort studies (RR 1.09, 95% CI 0.90 to 1.31, n = 3,110; Figure 49). The results were statistically homogeneous across all studies. One RCT[50] and one cohort study[95] reported emergency cesarean deliveries and found no difference (RCT, RR 0.81, 95% CI 0.62 to 1.05, n = 1,000; cohort, RR 0.83, 95% CI 0.33 to 2.06, n = 126).

Figure 49. Effect of treatment on outcomes of women with GDM: cesarean delivery

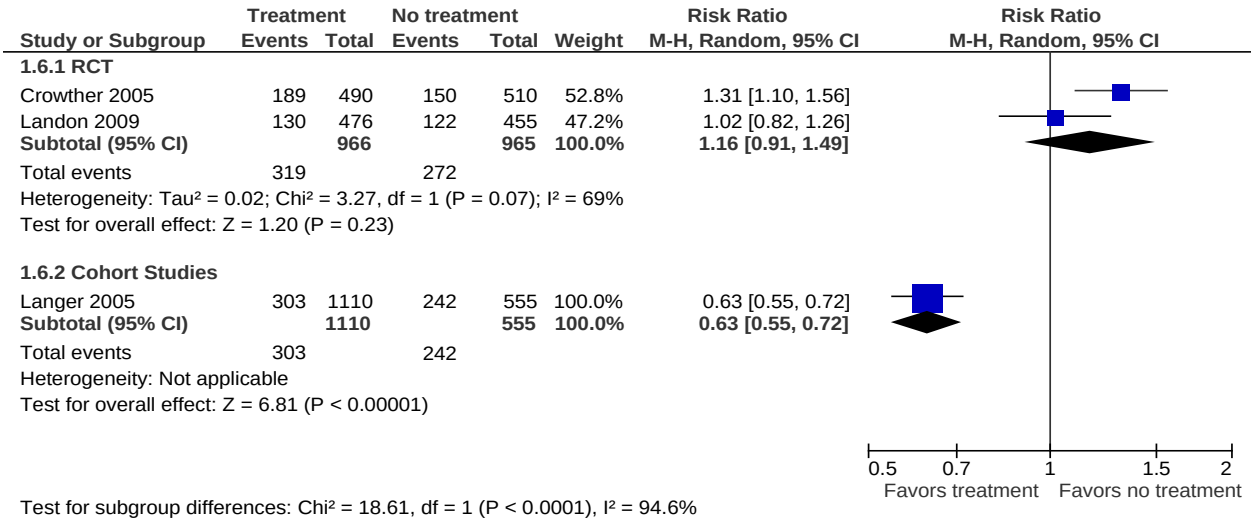


CI = confidence interval; GDM = gestational diabetes mellitus; RCT = randomized controlled trial; M-H = Mantel-Haenszel

Induction of Labor

Three studies provided data on induction of labor[50,54,146] but results differed significantly across the studies (Table 16). Two RCTs showed no significant difference overall (RR 1.16, 95% CI 0.91 to 1.49, n = 1,931), although there was important statistical heterogeneity between studies (I^2 = 69%). One RCT showed a significant difference favoring no treatment,[50] while the other study showed no difference (Figure 50).[54] Different study protocols may account for the heterogeneity of results between studies. In the study that showed more inductions of labor in the treatment group, no recommendations were provided regarding obstetrical care, thus replicating usual clinical care of women with GDM. In the other study, antenatal surveillance was reserved for standard obstetrical indications. In contrast the one cohort study showed a significant difference with fewer inductions in the treatment group (RR 0.63, 95% CI 0.55 to 0.72, n = 1,665).[146] Baseline differences in the study populations and regional practices may have accounted for the different results between studies. Further, the comparison group in the cohort study was women who presented late for obstetrical care which confounds the relationship between induction and GDM treatment. Furthermore, the cohort study protocol was to deliver these women within one week of GDM diagnosis so the outcome of induction was substantially confounded by different delivery protocols between treatment and nontreatment groups.

Figure 50. Effect of treatment on outcomes of women with GDM: induction of labor

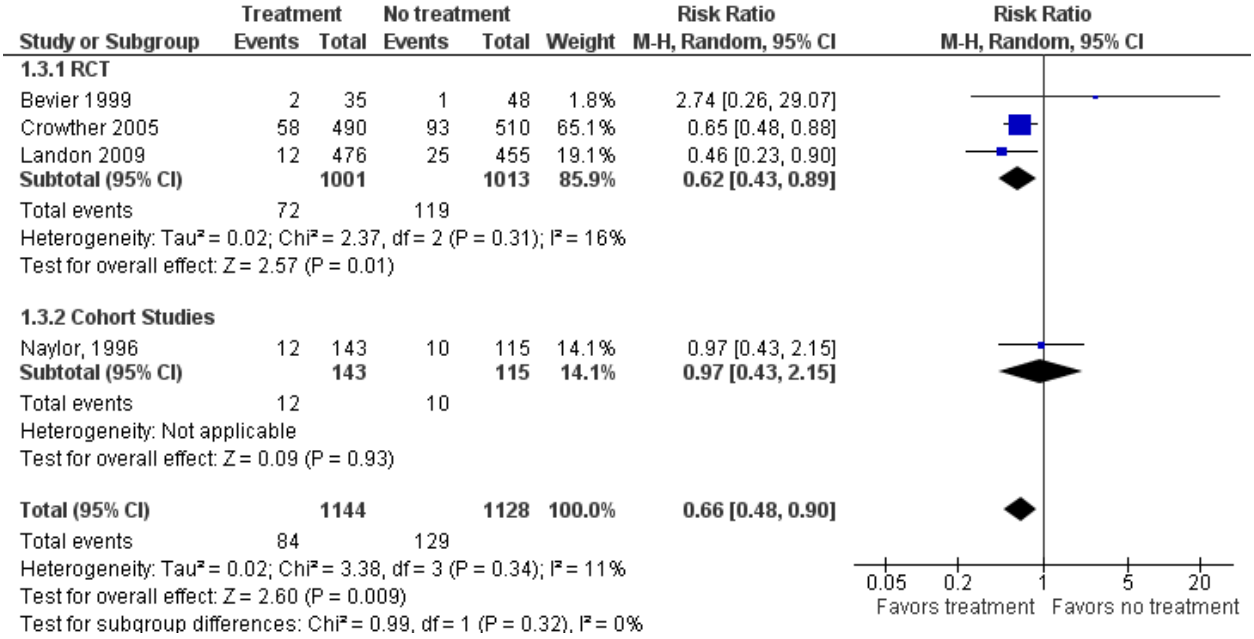


CI = confidence interval; GDM = gestational diabetes mellitus; RCT = randomized controlled trial; M-H = Mantel-Haenszel

Preeclampsia

Three RCTs and one cohort study provided data on preeclampsia (Table 16).[50,54,98,160] Pooled estimate for the RCTs showed a significant difference favoring the treated group (RR 0.62; 95% CI, 0.43 to 0.89, n = 2,014) with minimal statistical heterogeneity across studies (I^2 = 16%; Figure 51). The strength of evidence was considered moderate (Table 18). One of the studies also reported preeclampsia or gestational hypertension as a combined outcome,[54] and also showed a significant difference favoring the treatment group (RR 0.63; 95% CI, 0.44 to 0.92, n = 931). In all three trials, there was no significant difference between groups in gestational age at delivery.

Figure 51. Effect of treatment on outcomes of women with GDM: preeclampsia



CI = confidence interval; GDM = gestational diabetes mellitus; RCT = randomized controlled trial; M-H = Mantel-Haenszel

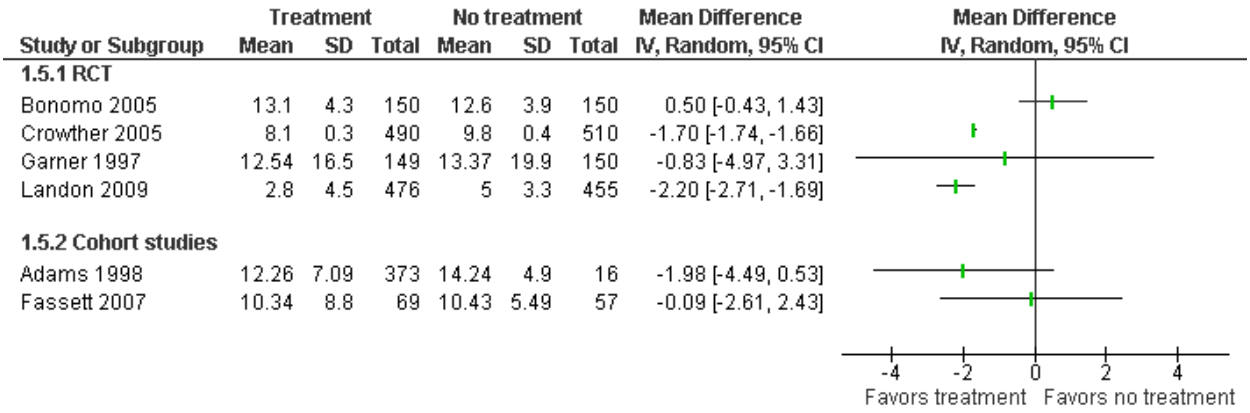
Birth Trauma

One study provided data on maternal birth trauma (postpartum hemorrhage).[92] No significant difference was observed between groups (Table 16).

Weight Gain

Six studies provided data on weight gain (Table 16).[50,54,95-97,152] Pooled results for the RCTs are not presented due to substantial heterogeneity (I^2=88%). Two RCTs showed no significant difference,[96,97] while two large RCTs showed a significant difference with less weight gain in the treatment group (Figure 52).[50,54] Given the high BMIs of the women studied in these large RCTs, less gestational weight gain in the treatment group could be interpreted as a beneficial finding. Pooled results for the cohort studies showed no significant difference between groups (mean difference [MD] -1.04; 95% CI, -2.89 to 0.81, n = 515). The strength of evidence was considered insufficient for this outcome (Table 18).

Figure 52. Effect of treatment on outcomes of women with GDM: weight gain



CI = confidence interval; GDM = gestational diabetes mellitus; IV = inverse variance; RCT = randomized controlled trial; SD = standard deviation

BMI at Delivery

Only one RCT reported BMI at delivery and showed a significantly lower BMI in the treated group compared with the untreated group (mean BMI 31.3 vs. 32.3; mean difference -1.00; 95% CI, -1.67 to -0.33, n = 931) (Table 16).[54]

Table 16. Evidence summary for Key Question 4: maternal outcomes

Outcome	Source	Number of Studies	Number of Participants	Effect Estimate*	I^2 (%)	Favors
Cesarean section	RCT	5	2613	0.90 [0.79, 1.01]	0	-
	Cohort	6	3110	1.09 [0.90, 1.31]	23	-
Unplanned cesarean section	RCT	1	1000	0.81 [0.62, 1.05]	NA	-
	Cohort	1	126	0.83 [0.33, 2.06]	NA	-
Induction of labor	RCT	2	1931	1.16 [0.91, 1.49]	69	-
	Cohort	1	1665	0.63 [0.55, 0.72]	NA	Treatment
Preeclampsia	RCT	3	2014	0.62 [0.43, 0.89]	16	Treatment
	Cohort	1	258	0.97 [0.43, 2.15]	NA	-
Preeclampsia or gestational hypertension	RCT	1	931	0.63 [0.44, 0.92]	NA	Treatment
	Cohort	1	874	0.30 [0.15, 0.62]	NA	-
Weight gain (kg)	RCT	4	2530	Pooled estimate not reported due to heterogeneity	88	-
	Cohort	2	515	-1.04 [-2.89, 0.81]†	8	-
Maternal birth trauma	Cohort	1	874	0.95 [0.21, 4.28]	NA	-
BMI at delivery	RCT	1	931	-1.00 [-1.67, -0.33]†	NA	Treatment

NA = not applicable; RCT = randomized controlled trial
*Risk ratios unless otherwise specified.
†Mean difference.

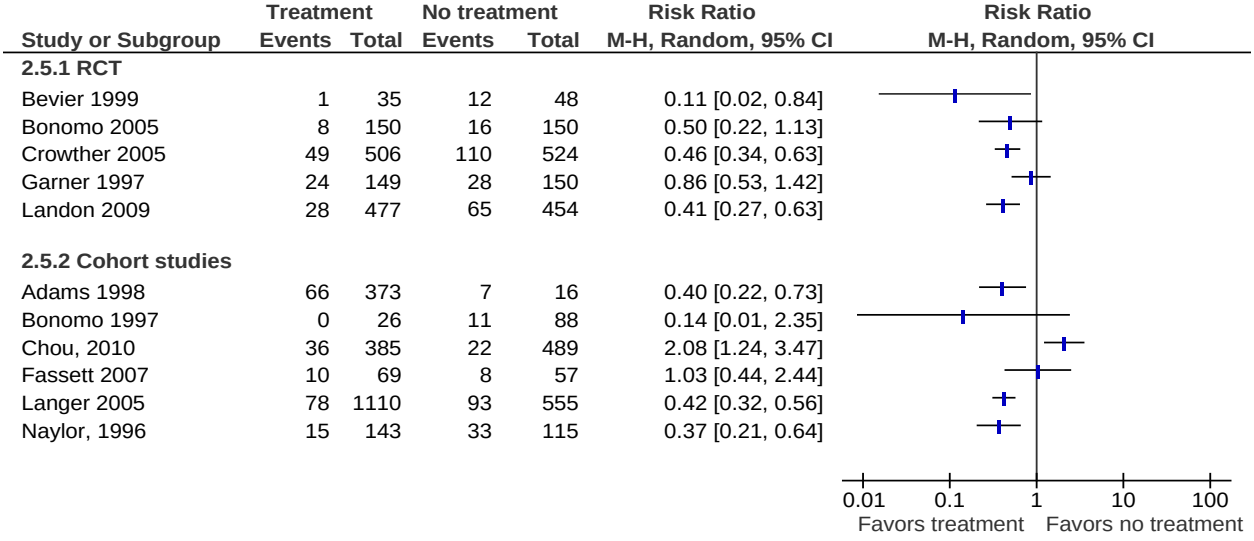
Fetal/Neonatal/Child Outcomes

Short Term

Birthweight

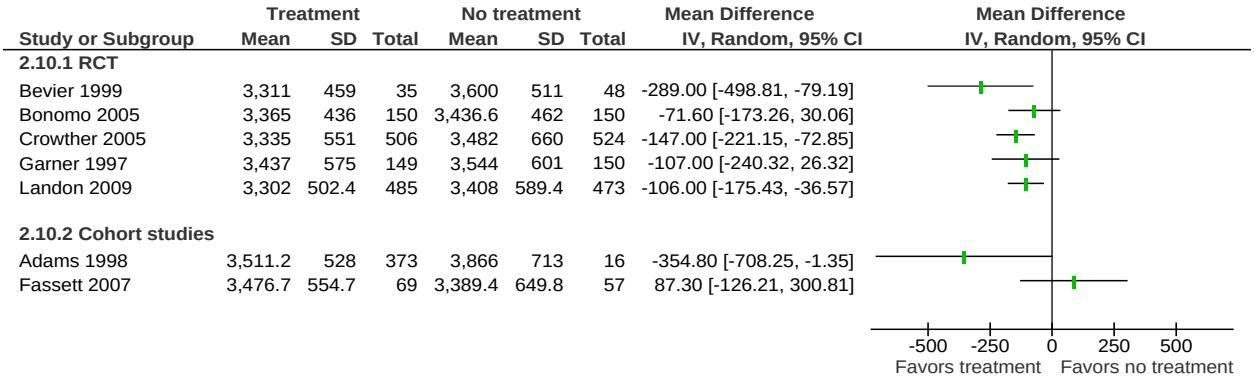
All studies reported birthweights for the infants (Table 17).[50,54,92,95-98,146,148,152,160] Pooled estimate for the RCTs showed significantly lower incidence of birthweights >4,000 g among infants in the treated groups (RR 0.50, 95% CI, 0.35 to 0.71; Figure 53); however, there was moderate heterogeneity across studies. Pooled estimates were not reported for the cohort studies because of substantial heterogeneity (I^2=86%). Three of the studies[96,152,160] also reported the incidence of birthweights >4,500 g and showed no significant differences between groups. In terms of actual birthweight (Figure 54), the five RCTs showed significantly lower mean birthweights among the treated group (MD -120.8; 95% CI, -163.4 to -78.2, n = 2,670). The two cohort studies showed substantial heterogeneity with one study showing a significantly lower mean birthweight in the treated group and the second cohort study showing no difference between groups.

Figure 53. Effect of treatment on outcomes for offspring of women with GDM: birthweight >4,000 g



CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; RCT = randomized controlled trial

Figure 54. Effect of treatment on outcomes for offspring of women with GDM: birthweight (continuous)

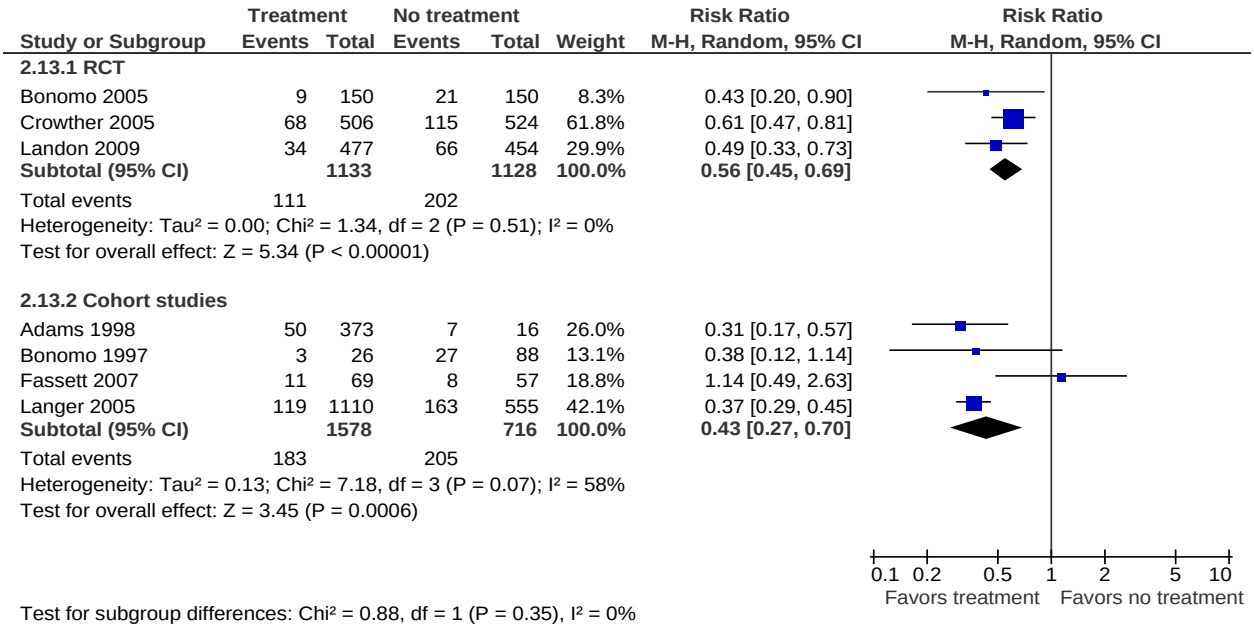


CI = confidence interval; GDM = gestational diabetes mellitus; IV = inverse variance; RCT = randomized controlled trial; SD = standard deviation

Large for Gestational Age (LGA)

There was a significant difference in LGA with the treatment group having fewer cases among both the three RCTs[50,54,97] (RR 0.56; 95% CI, 0.45 to 0.69, n = 2,261) and the four cohort studies (RR 0.43; 95% CI, 0.27 to 0.70, n = 2,294) (Table 17).[95,148,152,152] The results for the cohort studies showed moderate statistical heterogeneity (I^2 = 58%) (Figure 55). One study appeared to be an outlier,[95] and when removed from the analysis there was no heterogeneity.

Figure 55. Effect of treatment on outcomes for offspring of women with GDM: large for gestational age (LGA)

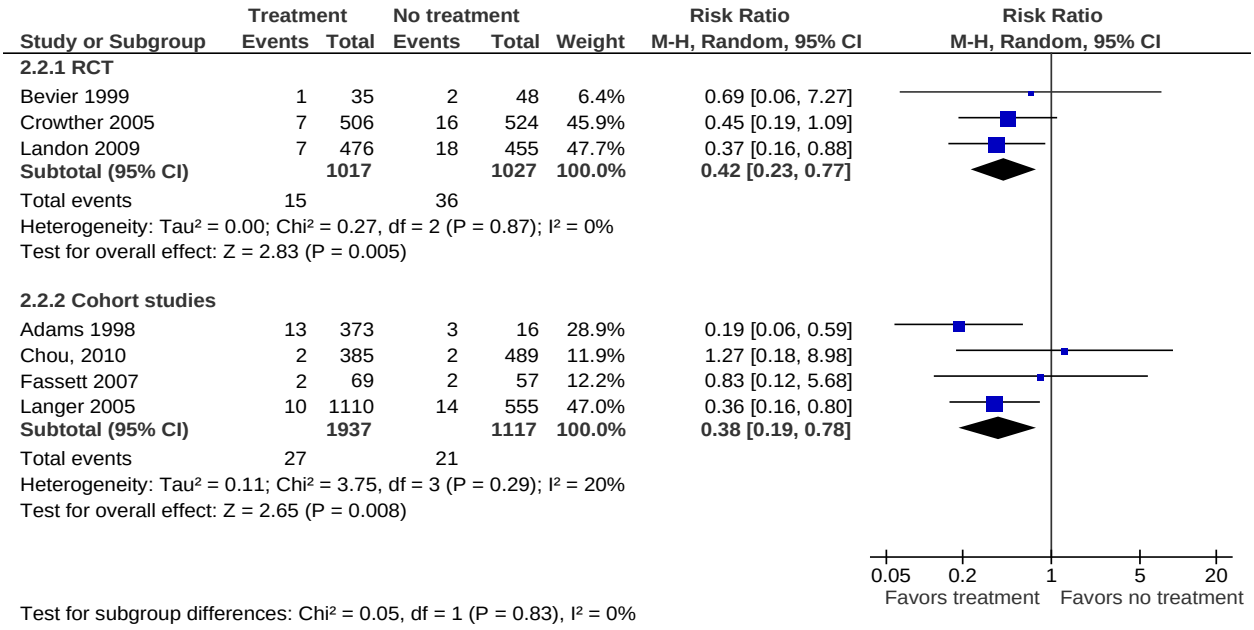


CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; RCT = randomized controlled trial

Shoulder Dystocia

Seven studies provided data on shoulder dystocia (Table 17).[50,54,92,95,98,146,152] Pooled estimates from three RCTs[50,54,98] showed a significant difference favoring the treated group (RR 0.42; 95% CI, 0.23 to 0.77, n = 2,044; (Figure 56). The four cohort studies[92,95,146,152] also showed a significant difference favoring the treated group (RR 0.38; 95% CI, 0.19 to 0.78, n = 3,054). There was no statistical heterogeneity across studies. Overall, the strength of evidence was considered moderate showing a difference in favor of the treated group. Shoulder dystocia was reduced for all studies combined; however, individual studies that included women with milder forms of glucose intolerance (i.e., OGCT screen positive OGTT negative, one RCT [98] and one cohort study[95]) showed no differences.

Figure 56. Effect of treatment on outcomes for offspring of women with GDM: shoulder dystocia



CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; RCT = randomized controlled trial

Brachial Plexus Injury

One RCT[50] and one cohort study[152] provided data for brachial plexus injury (Table 17). The RCT found no significant difference between groups (RR 0.15; 95% CI, 0.01 to 2.87, n = 1,000), while the cohort study showed a significant difference favoring the treated group (RR 0.04; 95% CI, 0 to 0.66, n = 389).

Clavicular Fracture

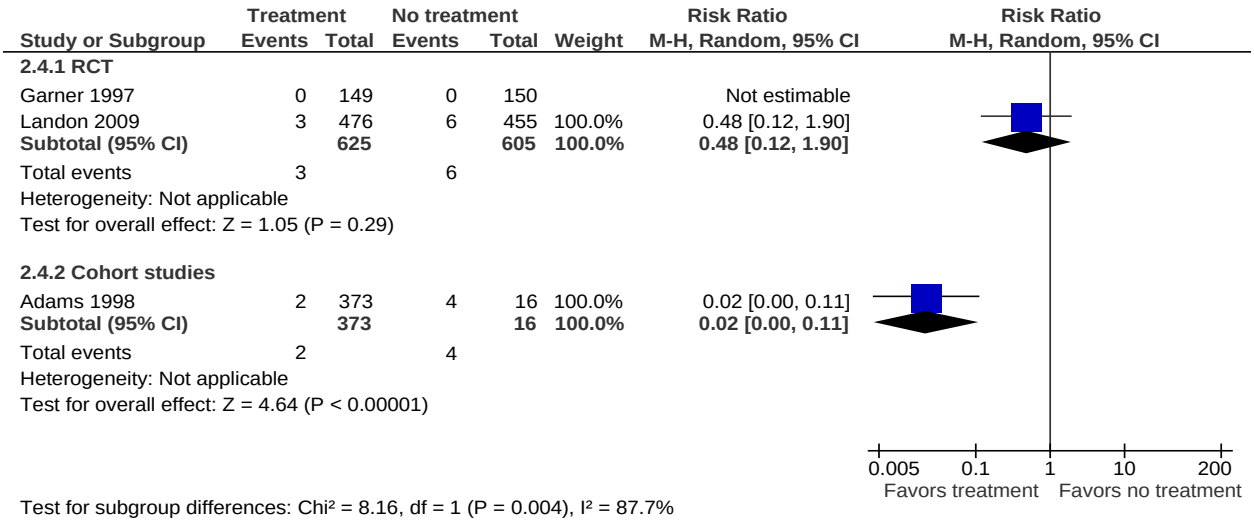
The same two studies[50,152] reported clavicular fractures with no difference for the RCT [50] (RR 0.35; 95% CI, 0.01 to 8.45, n = 1,030), and a significant difference favoring the treated group in the cohort study[152] (RR 0.02; 95% CI, 0 to 0.22, n = 389; Table 17).

Birth Trauma

Three studies reported birth trauma.[54,96,152] Birth trauma was defined as brachial plexus palsy or clavicular, humeral, or skull fracture in one study.[54] Brachial plexus injury, cranial nerve palsy, and clavicular facture were components of birth trauma in one study.[152] In the third study

birth trauma or injury included fractures and neurologic sequelae. One of the RCTs found no incidents in either group;[96] the second RCT[54] showed no significant difference between groups (RR 0.48; 95% CI, 0.12 to 1.90, n = 1,230; Figure 57). One cohort study showed a significant difference favoring the treated group (RR 0.02; 95% CI, 0.00 to 0.11, n = 389) (Table 17).[152] Overall, the strength of evidence was insufficient for this outcome (Table 18).

Figure 57. Effect of treatment on outcomes for offspring of women with GDM: birth trauma

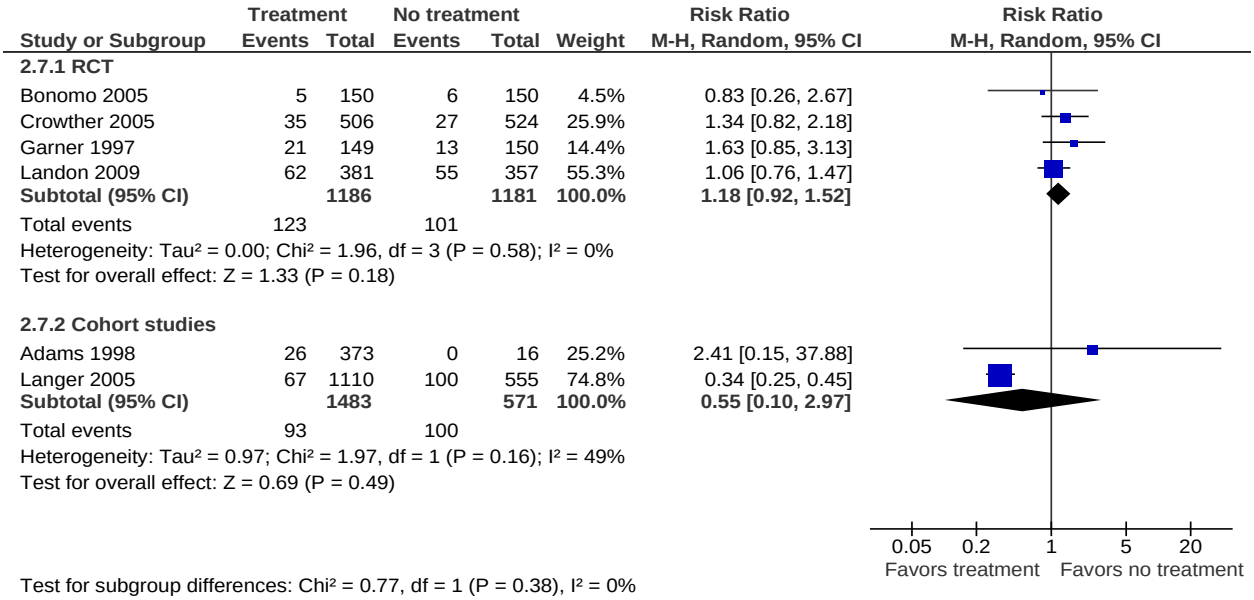


CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; RCT = randomized controlled trial

Hypoglycemia

Six studies provided data on hypoglycemia (Table 14).[50,54,96,97,146,152] The pooled results from four RCTs showed no significant difference between groups (RR 1.18; 95% CI, 0.92 to 1.52, n = 2,367) and no statistical heterogeneity (Figure 58). The two cohort studies showed different results: one study showed no significant difference, while the second study showed a significant difference favoring the treated group (overall RR 0.55; 95% CI, 0.10 to 2.97, n = 2,054). The different results may be due in part to different definitions of hypoglycemia used across the studies. Overall, the strength of evidence was low suggesting no difference between groups in the incidence of hypoglycemia (Table 15).

Figure 58. Effect of treatment on outcomes for offspring of women with GDM: hypoglycemia

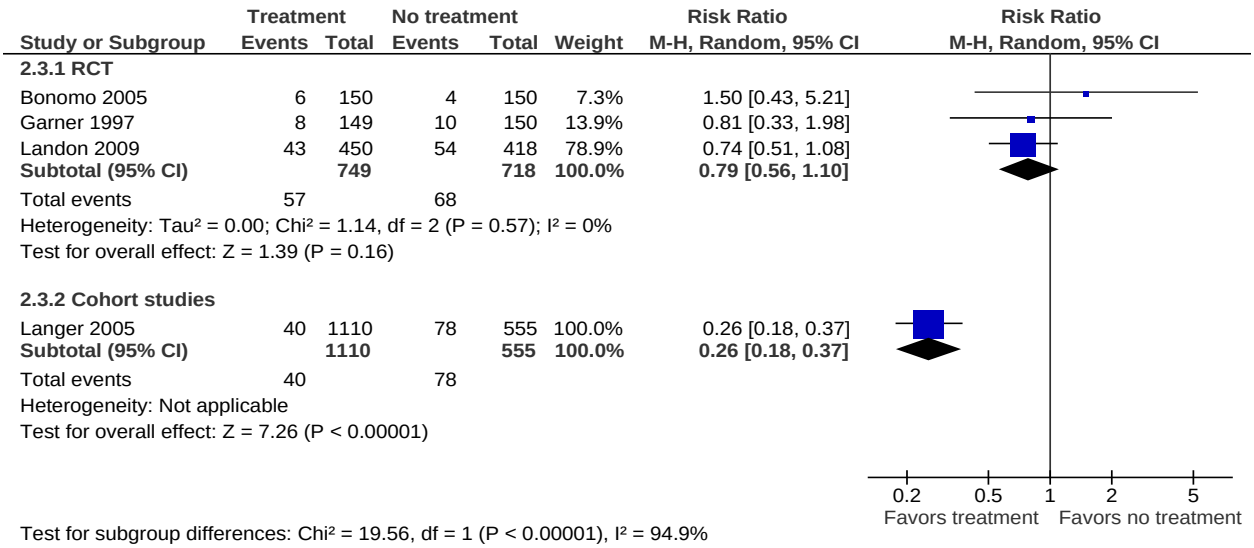


CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; RCT = randomized controlled trial

Hyperbilirubinemia

Four studies provided data on hyperbilirubinemia (Table 14).[54,96,97,146] Three RCTs showed no significant difference between groups[54,96,97] (RR 0.79; 95% CI, 0.56 to 1.10, n = 1,467), while one cohort study showed a significant difference favoring the treated group[146] (RR 0.26; 95% CI, 0.18 to 0.37, n = 1,665; Figure 59).

Figure 59. Effect of treatment on outcomes for offspring of women with GDM: hyperbilirubinemia

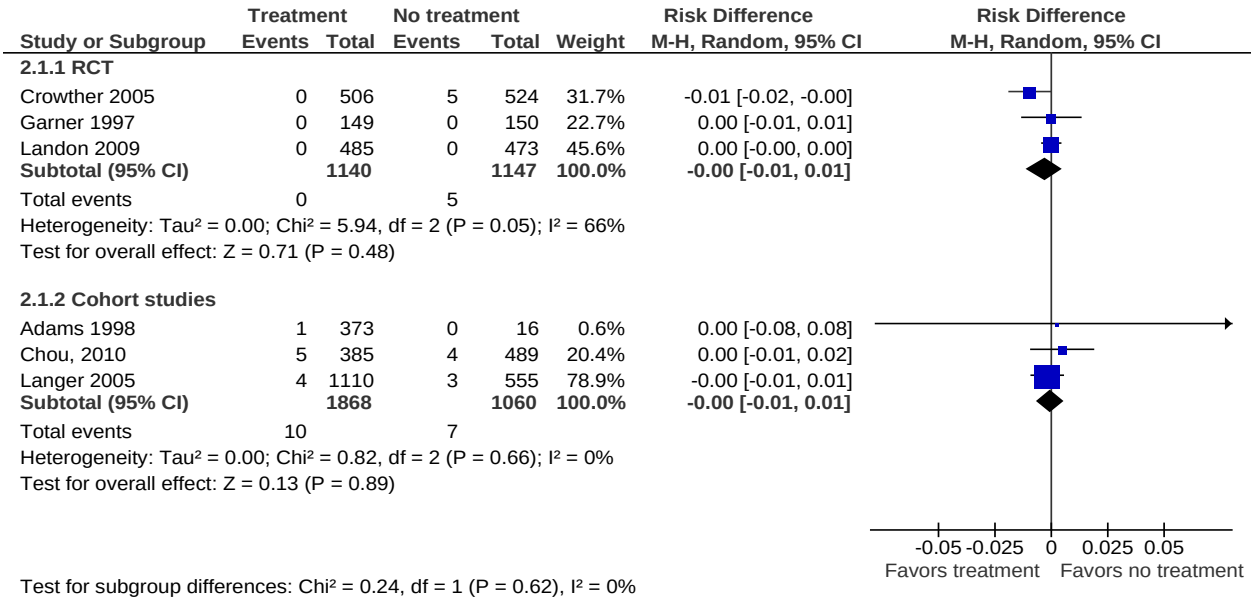


CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; RCT = randomized controlled trial

Mortality

Six studies provided data on perinatal deaths (Table 14).[50,54,92,96,146,152] No significant differences were found between groups for the three RCTs[50,54,96] (RD 0; 95% CI, -0.01 to 0.01, n = 2,287) or for the three cohort studies[92,146,152] (RD 0; 95% CI, -0.01 to 0.01, n = 2,928; Figure 60). There was heterogeneity among the three RCTs with one study showing a significant difference in favor of the treatment group.

Figure 60. Effect of treatment on outcomes for offspring of women with GDM: perinatal deaths

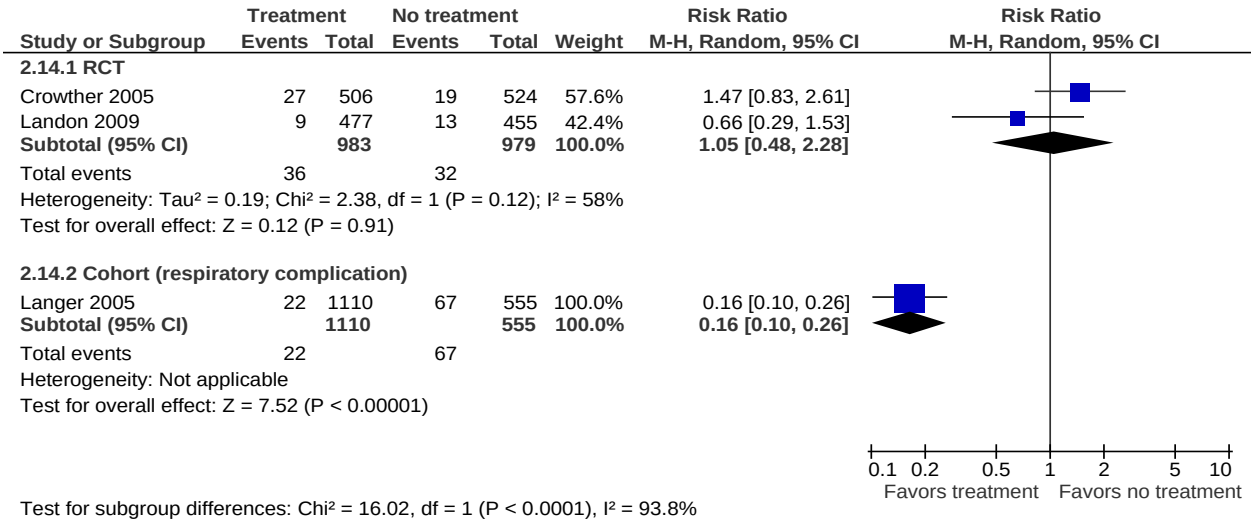


CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; RCT = randomized controlled trial

Respiratory Complications

Two RCTs[50,54] reported on respiratory distress syndrome and showed no significant difference between groups (RR 1.05; 95% CI, 0.48 to 2.28, n = 1,962; Table 17, Figure 61). One cohort study[146] reported respiratory complications and showed a significant difference favoring the treated group (RR 0.16; 95% CI, 0.10 to 0.26, n = 1,665).

Figure 61. Effect of treatment on outcomes for offspring of women with GDM: respiratory complications

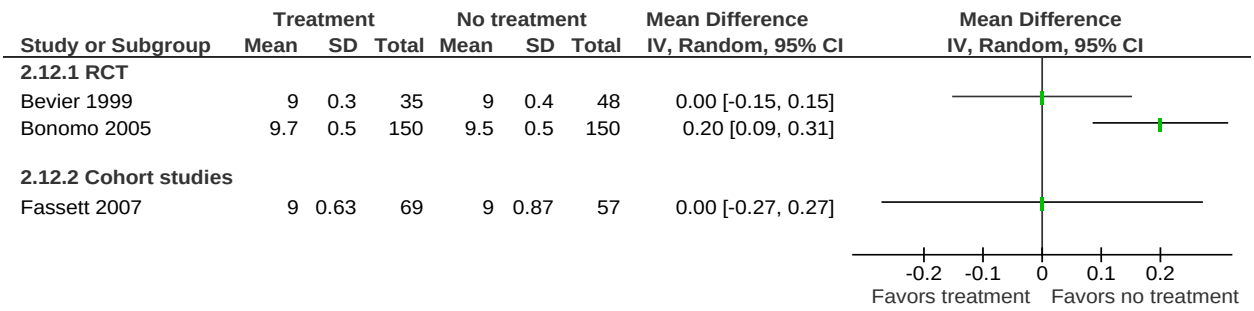


CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; RCT = randomized controlled trial

APGAR

One RCT[50] and one cohort study[95] compared APGAR scores at 1 minute (Table 17). Both showed a significant difference favoring the treatment group, although the results were more dramatic for the cohort study (RCT MD -0.30; 95% CI, -0.56 to -0.04, n = 83; cohort MD -1.00; 95% CI, -1.54 to -0.46, n = 126; Figure 56). Another cohort study[92] reported the number of infants with APGAR scores <7 at 1 minute and showed no difference between groups (RR 0.76, 95% CI, 0.47 to 1.25). Two RCTs[97,98] and one cohort study[95] compared APGAR scores at 5 minutes and no overall differences were found (Figure 62). There was substantial statistical heterogeneity between the two RCTs with one RCT showing no difference and the second showing a significant difference favoring the untreated group. The cohort study showed no difference (n = 126). One study[50] reported APGAR scores <7 at 5 minutes and found no difference between groups (n = 1,030).

Figure 62. Effect of treatment on outcomes for offspring of women with GDM: APGAR scores, 5 minutes



CI = confidence interval; GDM = gestational diabetes mellitus; IV = inverse variance; RCT = randomized controlled trial; SD = standard deviation

Other Infant Outcomes

Single studies reported on elevated cord blood c-peptide level,[54] preterm delivery,[54] length,[97] ponderal index,[97] any serious perinatal complication,[50] and abnormal fetal heart rate.[98] Significant differences were found for ponderal index (MD -0.09; 95% CI, -0.16 to -0.02, n = 300) and any serious perinatal complication (RR 0.32; 95% CI, 0.14 to 0.73, n = 1,030). Both results favored the treated group (Table 17).

Long Term

Type 2 Diabetes Mellitus

One small study reported 7 to 11 year followup and showed no significant difference in the incidence of type 2 diabetes mellitus among the offspring (RR 1.88; 95% CI, 0.08 to 44.76, n = 89).[96] The same study reported impaired glucose tolerance at 7-11 year followup.[96] Overall no difference was found (Table 17) (RR 5.63; 95% CI, 0.31 to 101.32, n = 89). The strength of evidence for both type 2 diabetes mellitus and impaired glucose tolerance was considered insufficient (Table 18).

BMI

One small study reported the incidence of BMI >95 percentile at 7 to 11 year followup and showed no significant difference between groups (RR 1.58; 95% CI, 0.66 to 3.79, n = 85; Table 17).[96] The original RCT[96] showed no differences in mean birth weight or macrosomia (birthweight >4,000 g and birthweight >4,500 g). A followup study[9] reporting outcomes at 4 to 5 years following the initial RCT reported BMI >85 percentile and also found no difference between groups (RR 1.19; 95% CI, 0.78 to 1.82, n = 199), despite a clear difference in macrosomia rates between treatment and control group (5% vs. 22%, respectively). When the two studies were pooled, the results showed no difference (RR 1.26; 95% CI, 0.86, 1.84, n = 284, Table 17) and the strength of evidence was considered low (Table 18).

Table 17. Evidence summary for Key Question 4: infant outcomes

Outcome	Source	Number of Studies	Number of Participants	Effect Estimate*	I^2 (%)	Favors
Birthweight >4,000 g	RCT	5	2,643	0.50 [0.35, 0.71]	50	Treatment
	Cohort	6	3,426	Results not pooled due to substantial heterogeneity	86	Treatment
Birthweight >4,500 g	RCT	1	299	1.01 [0.33, 3.05]	NA	-
	Cohort	2	647	0.29 [0.07, 1.25]	69	-
Birthweight	RCT	5	2,670	-120.81 [-163.40, -78.23]†	2	Treatment
	Cohort	2	515	Results not pooled due to substantial heterogeneity	77	-
Large for gestational age	RCT	3	2,261	0.56 [0.45, 0.69]	0	Treatment
	Cohort	4	2,294	0.43 [0.27, 0.70]	58	Treatment
Shoulder dystocia	RCT	3	2,044	0.42 [0.23, 0.77]	0	Treatment
	Cohort	4	3,054	0.38 [0.19, 0.78]	0	Treatment
Brachial plexus injury	RCT	1	1,000	0.15 [0.01, 2.87]	NA	-
	Cohort	1	389	0.04 [0.00, 0.66]	NA	Treatment
Clavicular fracture	RCT	1	1,030	0.35 [0.01, 8.45]	NA	-
	Cohort	1	389	0.02 [0.00, 0.22]	NA	Treatment
Birth trauma	RCT	2	1,230	0.48 [0.12, 1.90]	NA	-
	Cohort	1	389	0.02 [0.00, 0.11]	NA	Treatment
Hypoglycemia	RCT	4	2,367	1.18 [0.92, 1.52]	0	-
	Cohort	2	2,054	0.55 [0.10, 2.97]	49	-
Hyperbilirubinemia	RCT	3	1,467	0.79 [0.56, 1.10]	0	-
	Cohort	1	1,665	0.26 [0.18, 0.37]	NA	Treatment
Perinatal deaths	RCT	3	2,287	-0.00 [-0.01, 0.01]‡	66	-
	Cohort	3	2,928	-0.00 [-0.01, 0.01]‡	0	-
Respiratory complications	RCT (RDS)	2	1,962	1.05 [0.48, 2.28]	58	-
	Cohort (complications)	1	1,665	0.16 [0.10, 0.26]	NA	Treatment
APGAR 1 min	RCT	1	83	-0.30 [-0.56, -0.04]	NA	Treatment
	Cohort	1	126	-1.00 [-1.54, -0.46]	NA	Treatment
APGAR 5 min	RCT	2	383	Results not pooled due to substantial heterogeneity	77	-
	Cohort	1	126	0.00 [-0.27, 0.27]	NA	-

Table 17. Evidence summary for Key Question 4: infant outcomes (continued)

Outcome	Source	Number of Studies	Number of Participants	Effect Estimate*	I^2 (%)	Favors
Type 2 DM (long-term)	RCT	1	89	1.88 [0.08, 44.76]	NA	-
Impaired glucose tolerance	RCT	1	89	5.63 [0.31, 101.32]	44	-
BMI (long-term)	>95 percentile	1	85	1.58 [0.66, 3.79]	NA	-
	>85 percentile	1	199	1.19 [0.78, 1.82]	NA	-
	Any BMI (2 studies above combined)	2	284	1.26 [0.86, 1.84]	0	-

BMI = body mass index; DM = diabetes mellitus; NA = not applicable; RCT = randomized controlled trial; RDS = respiratory distress syndrome
*Risk ratios unless otherwise specified.
†Mean difference.
‡Risk difference.

Table 18. Strength of evidence for Key Question 4: maternal and infant outcomes

Outcome	Source	Risk of Bias	Consistency	Directness	Precision	Overall SOE	Comment
Preeclampsia	3 RCTs	Low	Consistent	Direct	Imprecise	Moderate (favors treatment)	The evidence provides moderate confidence that the estimate reflects the true effect in favor of the treatment group.
	1 cohort	High	Unknown	Direct	Imprecise	Insufficient	
Maternal weight gain	4 RCTs	Medium	Inconsistent	Direct	Imprecise	Insufficient	There is insufficient evidence to draw conclusions for this outcome
	2 cohorts	High	consistent	Direct	Imprecise	Insufficient	
Birth injury	2 RCTs	Medium	Consistent	Direct	Imprecise	Low	There is insufficient evidence to make a conclusion for this outcome. There is a difference in findings for the RCTs and cohort studies; the number of events and participants across all studies does not allow for a conclusion.
	1 cohort	High	Unknown	Direct	Imprecise	Insufficient (favors treatment)	
Shoulder dystocia	3 RCTs	Medium	Consistent	Direct	Precise	Moderate (favors treatment)	The evidence provides moderate confidence that the estimate reflects the true effect in favor of the treatment group.
	4 cohorts	High	Consistent	Direct	Precise	Low (favors treatment)	
Neonatal hypoglycemia	4 RCTs	Medium	Consistent	Direct	Imprecise	Low (no difference)	The evidence provides low confidence that there is no difference between groups.
	2 cohorts	High	Inconsistent	Direct	Imprecise	Insufficient	
Macrosomia >4,000 g	5 RCTs	Medium	Consistent	Direct	Precise	Moderate (favors treatment)	The evidence provides moderate confidence that the estimate reflects the true effect in favor of the treatment group.
	6 cohorts	High	Inconsistent	Direct	Precise	Low (favors treatment)	
Long-term metabolic outcomes: impaired glucose tolerance	1 RCT	Medium	Unknown	Direct	Imprecise	Insufficient	There is insufficient evidence to draw conclusions for this outcome.
Long-term metabolic outcomes: type 2 diabetes mellitus	1 RCT	Medium	Unknown	Direct	Imprecise	Insufficient	There is insufficient evidence to draw conclusions for this outcome.
Long-term metabolic outcomes: BMI (assessed as >85th and >95th percentile)	2 RCTs	Medium	Consistent	Direct	Imprecise	Low (no difference)	The evidence provides low confidence that there is no difference between groups.

BMI = body mass index; RCT = randomized controlled trial; SOE = strength of evidence

Key Question 5. What are the harms of treating GDM and do they vary by diagnostic approach?

Description of Included Studies

Five of the studies included in Key Question 4 also provided data for Key Question 5.[50,54,95,97,98] The studies are described in Appendix D. All studies compared diet modification, glucose monitoring and insulin as needed with standard care. Four of the studies were randomized controlled trials,[50,54,97,98] while one study was a retrospective cohort.[95] The studies were published between 1999 and 2009 (median year 2005). Two studies had two associated publications reporting initial and longer term outcomes.[163,164] Three studies were from the United States,[54,95,98] and one each from Italy[97] and Australia.[50] The screening test used in most studies was OGCT with a 100 g OGTT assessed using CC criteria, except for the study from Australia that used a OGCT with a 75 g OGTT. Timing of diagnosis of GDM occurred at or after 24 weeks' gestation. Among these studies a variety of glucose threshold criteria were used for inclusion, varying from 50 g screen positive with nondiagnostic oral glucose tolerance tests to WHO criteria for a diagnosis of GDM. The 2 largest RCTs by Crowther et al. and Landon et al.[50,54] used different glucose thresholds for entry in their trials: WHO and CC criteria with a fasting glucose <95 mg/dL (5.3 mmol/L), respectively. The mean fasting glucose levels at study entry were similar between these 2 trials.

Methodological Quality of Included Studies

Among the four RCTs, one had low[50] and three[54,97,98] had unclear risk of bias. The trials that were unclear most commonly did not report detailed methods for sequence generation and allocation concealment. Two trials[54,97] were unclear with respect to blinding of participants. One trial had incomplete reporting of outcome data.[98] The cohort study was high quality (7/9 points);[95] the primary limitation was not controlling for potential confounders.

Key Points

- There was no evidence for some of the outcomes stipulated in the protocol including costs and resource allocation. There was limited evidence for harms and the evidence related to anxiety and depression. There was also limited evidence for number of prenatal visits and admissions to NICU. Results are detailed below.

Maternal Outcomes

- A single RCT assessed depression and anxiety at 6 weeks after study entry and 3 months postpartum. There was no significant difference between groups in anxiety at either time point, although there were significantly lower rates of depression in the treatment group at 3 months postpartum.

Outcomes in the Offspring

- Four RCTs reported small for gestational age and found no significant difference.

Health System Outcomes

- Three RCTs and one cohort study provided data on admission to NICU and showed no significant differences overall. One trial was an outlier as it showed a significant

difference favoring the no treatment group. This difference may be attributable to site-specific policies and procedures.

- Two RCTs reported on the number of prenatal visits and generally found significantly more visits among the treatment groups. The same two studies showed a lower incidence of patients requiring insulin therapy in the untreated groups.
- There was inconsistency across studies in terms of induction of labor with no difference found for the 2 RCTs overall and a significant difference favoring the treatment group among the one cohort study included. Among the RCTs, one showed a significant difference with fewer cases in the group receiving no treatment,[50] while the other study showed no difference.[54] In the RCT that showed more inductions of labor in the treatment group, no recommendations were provided regarding obstetrical care, thus replicating usual clinical care of women with GDM. In the other RCT, antenatal surveillance was reserved for standard obstetrical indications.
- No differences between groups were found for cesarean section (5 RCTs, 6 cohorts) or unplanned cesarean section (1 RCT, 1 cohort).

Detailed Synthesis

Maternal Outcomes

Depression and Anxiety

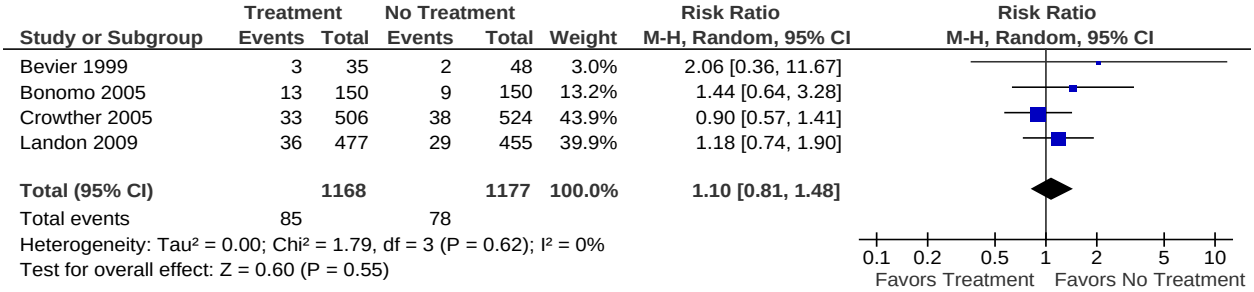
One RCT assessed depression and anxiety at 6 weeks after study entry and 3 months postpartum.[50] Depression was assessed using the Edinburgh Postnatal Depression Score and anxiety was assessed using the Spielberger State-Trait Anxiety Inventory. There was no significant difference between groups in anxiety at either time point, although there were significantly lower rates of depression in the treatment group 3 months postpartum (Table 19). The authors of the primary study note that the findings regarding anxiety and depression should be interpreted cautiously because they were based on a subgroup of the women included in the trial.

Fetal/Neonatal/Child Outcomes

Small for Gestational Age (SGA)

SGA was reported in four RCTs[50,54,97,98] and overall no significant difference was found between groups (RR 1.10; 95% CI, 0.81 to 1.48; Table 19, Figure 63).

Figure 63. Effect of treatment on adverse effects for infants of mothers with GDM: SGA



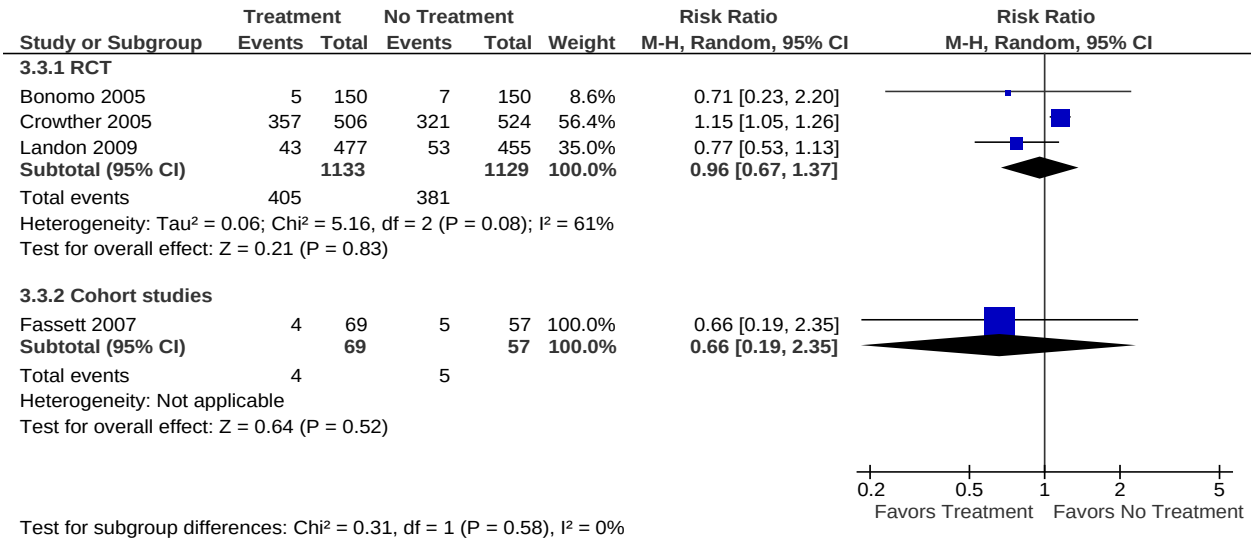
CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; SGA = small for gestational age

Society/Health Care System Outcomes

Admission to NICU

Three RCTs[50,54,97] and one cohort study[95] provided data on admission to the NICU (Table 19). Among the three RCTs there was no significant difference overall (RR 0.96; 95% CI, 0.67 to 1.37, n = 2,262; Table 19, Figure 64), although there was substantial statistical heterogeneity (I^2 = 61%). One study was an outlier as it showed a significant effect favoring the untreated group (RR 1.15; 95% CI, 1.05 to 126, n = 1,030). Removing this study from the analysis reduced the heterogeneity to 0% and the result remained non-significant. One cohort study also showed no significant difference in NICU admissions (RR 0.66; 95% CI, 0.19 to 2.35, n = 126).[95]

Figure 64. Effect of treatment on adverse effects for infants of mothers with GDM: NICU admissions



CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; NICU = neonatal intensive care unit; RCT = randomized controlled trial

Number of Prenatal Visits

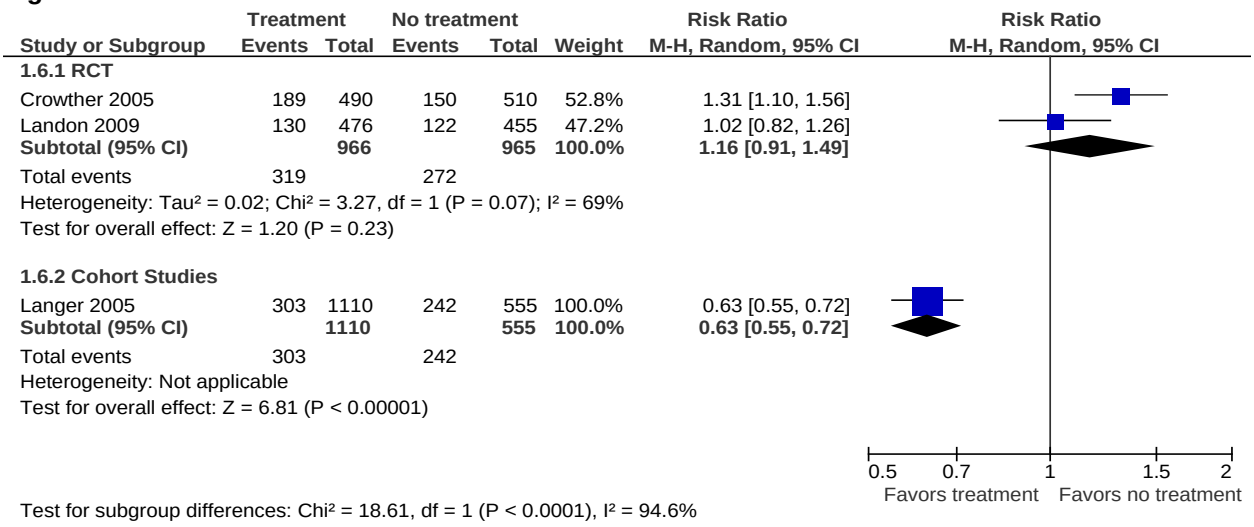
Two RCTs reported on the number of prenatal visits.[50,54] Landon et al.[54] reported an average of seven prenatal visits in the treatment group versus five in the control group (p<0.001). Crowther et al.[50] reported the median number of antenatal clinic visits and physician clinical visits after enrolment. The intervention group had fewer antenatal clinic visits (median 5.0 [inter-quartile range (IQR) 1-7] vs. 5.2 [IQR 3-7], p<0.001); whereas they had more physician clinic visits (median 3 [IQR 1-7] vs. 0 [IQR 0-2]). The intervention group also had significantly more visits with a dietician (92 percent vs. 10 percent, p<0.001) and with a diabetes educator (94 percent vs. 11 percent, p<0.001).

Induction of Labor

[Note: This outcome was presented under Key Question 4. It is also presented here as it may be considered a harm in terms of more resource use and more invasive management.] Three studies provided data on induction of labor[50,54,146] but results differed significantly across the studies (Table 19). Two RCTs showed no significant difference overall (RR 1.16, 95% CI 0.91 to 1.49, n = 1,931), although there was important statistical heterogeneity between studies (I^2 =

69%). One RCT showed a significant difference favoring no treatment,[50] while the other study showed no difference (Figure 65).[54] Different study protocols may account for the heterogeneity of results between studies. In the study that showed more inductions of labor in the treatment group, no recommendations were provided regarding obstetrical care, thus replicating usual clinical care of women with GDM. In the other study, antenatal surveillance was reserved for standard obstetrical indications. In contrast the one cohort study showed a significant difference with fewer inductions in the treatment group (RR 0.63, 95% CI 0.55 to 0.72, n = 1,665).[146] Baseline differences in the study populations and regional practices may have accounted for the different results between studies. Further, the comparison group in the cohort study was women who presented late for obstetrical care which confounds the relationship between induction and GDM treatment. Furthermore, the cohort study protocol was to deliver these women within one week of GDM diagnosis so the outcome of induction was substantially confounded by different delivery protocols between treatment and nontreatment groups.

Figure 65. Effect of treatment on outcomes of women with GDM: induction of labor

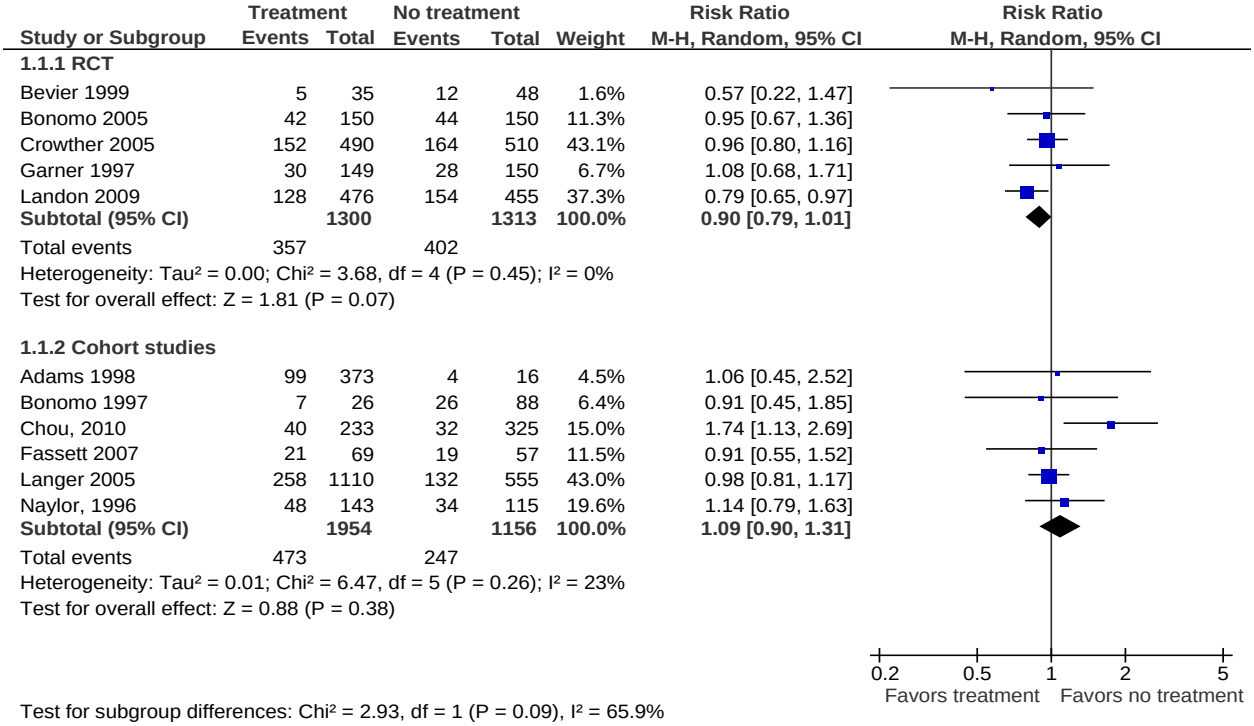


CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; RCT = randomized controlled trial

Cesarean Delivery

[Note: This outcome was presented under Key Question 4. It is also presented here as it may be considered a harm in terms of more resource use and more invasive management.] All studies provided data on cesarean delivery (Table 19).[50,54,92,95-98,146,148,152,160] There was no significant difference in the pooled estimates for the RCTs (RR 0.90, 95% CI 0.79 to 1.01, n = 2,613) or for the cohort studies (RR 1.09, 95% CI 0.90 to 1.31, n = 3,110; Figure 66). The results were statistically homogeneous across all studies. One RCT[50] and one cohort study[95] reported emergency cesarean deliveries and found no difference (RCT, RR 0.81, 95% CI 0.62 to 1.05, n = 1,000; cohort, RR 0.83, 95% CI 0.33 to 2.06, n = 126).

Figure 66. Effect of treatment on outcomes of women with GDM: cesarean delivery



CI = confidence interval; GDM = gestational diabetes mellitus; M-H = Mantel-Haenszel; RCT = randomized controlled trial

Table 19. Evidence summary for Key Question 5

Outcome	Number of Studies	Number of Participants	Effect Estimate*	I2 (%)	Favors
Small for gestational age (RCTs)	4	2,345	1.10 [0.81, 1.48]	0	-
Anxiety (6 weeks, RCT)	1	682	-0.30 [-0.88, 0.28]	NA	-
Anxiety (3 months, RCT)	1	573	-0.20 [-0.83, 0.43]	NA	-
Depression (3 months, RCT)	1	568	0.50 [0.31, 0.79]	NA	Treatment
Admission to NICU					
RCT	3	2,262	0.96 [0.67, 1.37]	61	-
Cohort	1	126	0.66 [0.19, 2.35]	NA	-
Induction of labor					
RCT	2	1,931	1.16 [0.91, 1.49]	69	-
Cohort	1	1,665	0.63 [0.55, 0.72]	NA	Treatment
Cesarean section					
RCT	5	2,613	0.90 [0.79, 1.01]	0	-
Cohort	6	3,110	1.09 [0.90, 1.31]	23	-
Unplanned cesarean section					
RCT	1	1000	0.81 [0.62, 1.05]	NA	-
Cohort	1	126	0.83 [0.33, 2.06]	NA	-

NA = not applicable; NICU = neonatal intensive care unit; RCT = randomized controlled trial
*****Risk ratio

Discussion

Key Findings and Discussion

Key findings are presented by Key Question in the sections that follow. A summary of the results for all Key Questions is provided in Table 24 at the end of the Discussion.

Key Question 1

Fifty-one studies provided data for Key Question 1 that sought to examine the test characteristics and prevalence of current screening and diagnostic tests for gestational diabetes mellitus (GDM). The lack of a "gold standard" to confirm a diagnosis of GDM limits the ability to compare the results of studies that used different diagnostic criteria. Different criteria result in different rates of prevalence for GDM, regardless of similarities across study settings and patient characteristics.

The methodological quality of the studies was assessed using the QUADAS-2 tool. There were several concerns about the quality and applicability of the studies that addressed Key Question 1. First, there is concern about the risk for partial verification bias, which can occur when not all of the patients are verified by the reference standard. In 25 percent of the studies, women who were below the threshold for further screening on the oral glucose challenge test (OGCT) did not undergo an oral glucose tolerance test (OGTT) to confirm a diagnosis of GDM. For 35 percent of studies, it was unclear risk whether all patients underwent both tests. Another concern relates to the risk of diagnostic review bias in which the interpretation of the results of the reference standard may have been influenced by the knowledge of the results of the index test. Eighty percent of studies were assessed as high or unclear risk for this domain. A third concern relates to the domain of patient selection and the possibility of spectrum bias; 82 percent of studies were assessed as having high or unclear concerns for applicability. This was primarily because the studies were conducted in developing countries and used the World Health Organization (WHO) criteria to diagnose GDM.

The evidence showed that the 50 g OGCT with the 130 mg/dL cutpoint had higher sensitivity when compared with the 140 mg/dL cutpoint; however, specificity was lower (99 vs. 85 and 77 vs. 86, respectively). Both thresholds have high negative predictive values (NPV), but variable positive predictive values (PPV) across a range of GDM prevalence. When the risk of a missing a diagnosis is considered high, screening tests with high NPV are preferred at the expense of PPV. However, if the harm of an incorrect diagnostic is high, screening tests with high PPV are preferred at the expense of NPV. The Toronto Trihospital study found evidence to support the use of the lower screening cutpoint for higher risk patients, and the higher screening cutpoint for lower risk patients.[15] While graded cutpoints for the diagnosis and treatment of dyslipidemia and osteoporosis based on risk factors are used in routine clinical practice, this approach is not widely accepted for the screening of GDM.

The large randomized controlled trials (RCTs) that showed some treatment benefits employed a two-step approach to screening and diagnosis for GDM.[50,54] The practical efficiency of a two-step approach may be improved by setting a high threshold value on the screening test, above which no further confirmation testing is required for diagnosis. One study provides support for this approach by demonstrating that a threshold of 200 mg/dL on a 50 g OGCT

resulted in 100 percent positive and negative predictive values for diagnosing GDM by Carpenter and Coustan (CC) and National Diabetes Data Group (NDDG) criteria.[104]

Only three studies included women who were in their first trimester of pregnancy and they used different diagnostic criteria. Therefore, no conclusions can be made about the test characteristics of screening tests for this group of women.

There are limited data to support the use of glycated hemoglobin (HbA1c) as a screening test. A study conducted in the United Arab Emirates using an HbA1c value of 5.5 percent or more lacked specificity (21 percent) despite good sensitivity (82 percent).[113] A study conducted in Turkey showed that an HbA1c cutoff of 7.2 percent or more had 64 percent sensitivity and specificity.[74] HbA1c does not perform as well as the 50 g OGCT as a screening test for GDM. However, when HbA1c is markedly elevated this supports a possible diagnosis of overt diabetes discovered in pregnancy. Since 2011-2012 the American Diabetes Association (ADA) has endorsed the use of an HbA1c of 6.5 percent or more as diagnostic of diabetes in nonpregnant women.[36] Studies of HbA1c with trimester specific cutoffs to determine the value at which overt diabetes should be diagnosed in pregnancy are needed.

The sensitivity for fasting plasma glucose (FPG) of 85 mg/dL as a screening test for GDM is similar to that for the 50 OGCT with a threshold of 140 mg/dL; however, specificity is lower. As the threshold for fasting glucose is increased specificity is gained at the expense of sensitivity. The use of fasting glucose as a screening test for GDM has several clinical advantages over the OGCT when the tests are performed at or after 24 weeks' gestation. FPG has the advantage of greater reproducibility than post glucose load testing.[165] In addition, it is easier to administer to women who cannot tolerate the glucose drink. Furthermore, fasting glucose has been positively associated with clinical outcomes of concern for GDM.[142,166] However, a recent report from the Hyperglycemia and Adverse Pregnancy Outcome Study (HAPO) data found that a fasting glucose of 92 mg/dL did not diagnose GDM in women from Hong Kong and Bangkok as frequently as it did in other populations, and the elevated post glucose load glucose measurements were more frequently diagnostic of an International Association of the Diabetes in Pregnancy Study Groups (IADPSG) diagnosis of GDM in women from Hong Kong and Bangkok.[6]

Our review did not identify compelling evidence for or against risk factor-based screening. Naylor et al. used the Toronto Trihospital study data to develop a risk scoring system for GDM screening using variable glucose thresholds based on age, body mass index (BMI), and race. When applied to a validation group, sensitivity and specificity were similar to universal screening.[167]

There are limited data to draw conclusions about the effectiveness of the different options for diagnostic testing for GDM. Four studies compared the 75 g and 100 g load tests, but they were conducted in different countries and used different criteria or thresholds. However, because both the 75 g and 100 g load tests are positively linked with outcomes[142,166] and the 75 g test is less time consuming, the adoption of the 75 g glucose load may be warranted even if thresholds continue to be debated[3,142]

The IADPSG has proposed the elimination of a screening test in favor of proceeding directly to a diagnostic test for GDM. We identified only one study[124] that compared the IADPSG criteria with the Australasian Diabetes in Pregnancy Society (ADIPS) that used a two-step strategy. Sensitivity was 82 percent (95% CI, 74 to 88) and specificity was 94 percent (95% CI, 93 to 96).

Prevalence and Predictive Values

The prevalence of GDM varied across studies and the diagnostic criteria used. Factors contributing to the variability included differences in study setting (i.e., country), screening practices (e.g., universal vs. selective), and population characteristics (e.g., race/ethnicity, age, BMI).

The predictive value of a screening or diagnostic test is determined by the test's sensitivity and specificity and by the prevalence of GDM. Table 20 presents a series of scenarios that demonstrate the changes in PPV and NPV for three levels of prevalence (7 percent, 15 percent, and 25 percent).[6] Separate tables are presented for different screening and diagnostic criteria. The higher the prevalence of GDM, the higher the PPV, or the more likely a positive result is able to predict the presence of GDM. When the prevalence of GDM is low, the PPV is also low, even when the test has high sensitivity and specificity. Generally the NPV (negative result rules out GDM) is very high—98 percent or better at a GDM prevalence of 7 percent.

Table 20. Relationship between predictive values and prevalence for different screening tests

Screening Test	Prevalence	Positive Predictive Value	Negative Predictive Value
50 g OGCT ≥140 mg/dL by CC/ADA (2000-2010) Sensitivity=85%; Specificity=86%	7%	31%	99%
	15%	52%	97%
	25%	67%	95%
50 g OGCT ≥130 mg/dL by CC/ADA (2000-2010) Sensitivity=99%; Specificity=77%	7%	24%	100%
	15%	43%	100%
	25%	59%	100%
50 g OGCT ≥140 mg/dL by NDDG Sensitivity=85%; Specificity=83%	7%	27%	99%
	15%	47%	97%
	25%	63%	94%
50 g OGCT ≥130 mg/dL by NDDG Sensitivity = 88%; Specificity = 66% (median)	7%	16%	99%
	15%	31%	97%
	25%	46%	94%
50 g OGCT ≥140 mg/dL by ADA 75 g Sensitivity=88%; Specificity=84% (median)	7%	29%	99%
	15%	49%	98%
	25%	65%	95%
50 g OGCT ≥140 mg/dL by WHO Sensitivity=78%; Specificity=81% (median)	7%	24%	98%
	15%	42%	95%
	25%	58%	92%
FPG (≥85 mg/dL) by CC/ADA (2000-2010) Sensitivity=87%; Specificity=52%	7%	12%	98%
	15%	24%	96%
	25%	38%	92%
Risk factor screening by various criteria Sensitivity=84%; Specificity=72% (median)	7%	21%	98%
	15%	38%	96%
	25%	54%	93%

TADA = American Diabetes Association; CC = Carpenter-Coustan; FPG = fasting plasma glucose; NDDG = National Diabetes Data Group; OGCT = oral glucose challenge test; WHO =World Health Organization

Key Question 2

Only two retrospective cohort studies were relevant to Key Question 2 which asked about the direct benefits and harms of screening for GDM. One retrospective cohort study (n=1,000) conducted in Thailand showed a significantly greater incidence of cesarean deliveries in the screened group. A survey of a subset of participants (n=93) in a large prospective cohort study involving 116,678 nurses aged 25-42 years in the United States found the incidence of macrosomia (infant weight ≥ 4.3 kg) was the same in the screened and unscreened groups (7 percent each group).

There were no RCTs available to answer questions about screening. There is a paucity of evidence on the impact of screening women for GDM on health outcomes. The comparison for this question was women who had and had not undergone screening. Since screening is now commonplace it may be unlikely to identify studies or cohorts where this comparison is feasible.

Key Question 3

Thirty-eight studies provided data for Key Question 3 that sought to examine health outcomes for women who meet various criteria for GDM and do not receive treatment. The majority of data came from cohort studies or the untreated groups from RCTs.

A wide variety of diagnostic criteria and thresholds were compared across the studies. The most common groups reported and compared were GDM diagnosed by CC criteria, no GDM by any criteria (normal), impaired glucose tolerance defined as one abnormal glucose value (OAV), and false positive (positive OGCT, negative OGTT).). Only single studies contributed data for many of the comparisons and outcomes, which does not allow for definitive conclusions. Further, results that showed no statistically significant differences cannot be interpreted as equivalence between groups nor do they rule out potential differences. A summary of the strength of evidence for key outcomes is provided in Table 21 and Table 22.

For maternal outcomes among the studies that compared groups as described above, women without GDM and those testing false positive showed fewer cases of preeclampsia than those meeting CC criteria; the strength of evidence was considered low for these two comparisons. No differences in preeclampsia were found for other comparisons, although evidence was based on few studies per comparison and strength of evidence was rated insufficient.

Fewer cases of cesarean section were found among women without GDM compared with women meeting criteria for CC GDM, CC, 1 abnormal OGTT, CC false positives, NDDG false positives, NDDG 1 abnormal oral glucose tolerance test, WHO IGT, IADPSG impaired fasting glucose (IFG), and IADPSG impaired glucose tolerance (IGT) IFG. There were fewer cases of cesarean section among false positives compared with women meeting criteria for CC GDM. For 12 other comparisons, there were no differences in rates of cesarean delivery.

For maternal hypertension, significant differences were found for eight of 16 comparisons; many comparisons were based on single studies. No GDM groups showed lower incidence of maternal hypertension when compared with CC GDM, CC, 1 abnormal OGTT, IADPSG IFG, IADPSG double impaired glucose tolerance (IGT-2), and IADPSG IGT IFG. Other comparisons showing significant differences were CC GDM versus false positives (lower incidence for false positives), IADPSG IGT versus IGT IFG (lower incidence for IGT), and IADPSG IFG versus IGT IFG (lower incidence for IFG).

Based on single studies, no differences were observed for maternal birth trauma for three comparisons. For maternal weight gain (less weight gain considered beneficial), significant differences were found for three of 12 comparisons: IADPSG IGT versus no GDM (favored IGT), IADPSG IFG versus no GDM (favored IFG), IADPSG IGT-2 versus no GDM (favored IGT-2). All comparisons were based on single studies and the strength of evidence was insufficient. For maternal mortality/morbidity, single studies contributed to three comparisons and no differences were found except for fewer cases among patient groups with no GDM compared with IADPSG GDM. No studies provided data on long-term maternal outcomes, such as type 2 diabetes mellitus, obesity and hypertension.

Table 21. Summary of strength of evidence for the association between different glucose levels and maternal outcomes (Key Question 3)

Outcome		Number of Studies	Strength of Evidence	Summary
Preeclampsia	CC GDM vs. no GDM	3 cohorts	Low	Statistically significant difference with fewer cases in the patient groups with no GDM (RR 1.50, 95% CI 1.07, 2.11)
	CC GDM vs. false positive	2 cohorts	Low	Statistically significant difference with fewer cases in the false-positive group (RR 1.51, 95% CI 1.17, 1.93)
	NDDG false positive vs. no GDM	2 cohorts	Insufficient	-
	NDDG, 1 abnormal OGTT vs. no GDM	1 cohort	Insufficient	-
	WHO IGT vs. no GDM	3 cohorts	Insufficient	-
Maternal weight gain	CC, 1 abnormal OGTT vs. no GDM	1 cohort	Insufficient	-
	WHO IGT vs. no GDM	1 cohort	Insufficient	-
	IADPSG IGT vs. no GDM	1 cohort	Insufficient	-
	IADPSG IFG vs. no GDM	1 cohort	Insufficient	-
	IADPSG IGT-2 vs. no GDM	1 cohort	Insufficient	-
	IADPSG IGT IFG vs. no GDM	1 cohort	Insufficient	-
	IADPSG IGT vs. IFG	1 cohort	Insufficient	-
	IADPSG IGT vs. IGT-2	1 cohort	Insufficient	-
	IADPSG IGT vs. IGT IFG	1 cohort	Insufficient	-
	IADPSG IFG vs. IGT-2	1 cohort	Insufficient	-
	IADPSG IFG vs. IGT IFG	1 cohort	Insufficient	-
	IADPSG IGT-2 vs. IGT IFG	1 cohort	Insufficient	-

CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; IFG = impaired fasting glucose; IGT = impaired glucose tolerance; IGT-2 = double impaired glucose tolerance; IADPSG = International Association of Diabetes and Pregnancy Study Groups; NDDG = National Diabetes Data Group; OGTT = oral glucose tolerance test; RR = risk ratio; WHO = World Health Organization

The most commonly reported outcome for the offspring was macrosomia >4,000 g. Six of 11 comparisons showed a significant difference: there were fewer cases in the group without GDM compared with CC GDM, CC 1 abnormal OGTT, NDDG GDM (unrecognized), NDDG false positives, and WHO IGT. Fewer cases were found for women with false-positive results compared with CC GDM. The strength of evidence for these findings was low to insufficient. Data for macrosomia >4,500 g was available for four comparisons and showed significant differences in two cases: patient groups with no GDM had fewer cases compared with women with CC GDM and with unrecognized NDDG GDM. The strength of evidence was low and insufficient, respectively.

For shoulder dystocia, significant differences were found for seven of 17 comparisons; all but one comparison was based on single studies (insufficient strength of evidence). Patient groups with no GDM showed lower incidence of shoulder dystocia when compared with CC GDM (5 studies; low strength of evidence), NDDG GDM (unrecognized), NDDG false positive, WHO IGT, IADPSG IFG, and IADPSG IGT IFG. The other significant difference showed lower incidence among the false-positive group compared with CC 1 abnormal OGTT. For fetal birth

trauma or injury, four studies compared CC GDM, NDDG GDM and WHO IGT with patient groups without GDM (insufficient strength of evidence). No differences were observed except for NDDG GDM which favored the group with no GDM.

Only one difference was found for neonatal hypoglycemia with fewer cases among patient groups without GDM compared with those meeting CC criteria; strength of evidence was insufficient. There were 16 comparisons for hyperbilirubinemia; the majority were based on single studies. Three comparisons showed significant differences between groups: patient groups with no GDM had fewer cases compared with CC false positive, IADPSG IGT, and IADPSG IGT-2, respectively. No differences were found for fetal morbidity/mortality for any of eight comparisons which may be attributable to small numbers of events within some comparisons. Moreover, comparisons were based on single studies.

Based on a single study, significant differences were found in prevalence of childhood obesity for CC GDM versus no GDM (lower prevalence for no GDM) and CC GDM versus false positives (lower prevalence for false positives). This was consistent for both childhood obesity >85th percentile as well as >95th percentile. However, this study was unable to control for maternal weight or BMI which are established predictors of childhood obesity. No differences, based on the same single study, were found for the other four comparisons within >85th or >95th percentiles. No other studies provided data on long-term outcomes, including type 2 diabetes mellitus and transgenerational GDM.

In summary, different thresholds of glucose intolerance impact maternal and neonatal outcomes of varying clinical importance. While many studies have attempted to measure the association between various criteria for GDM and pregnancy outcomes in the absence of treatment, the ability of a study or pooled analysis to find a statistically significant difference in pregnancy outcomes appears more dependent on study design, in particular the size of the study or pooled analysis, rather than the criteria used for diagnosing GDM. This is not surprising given the strong support found for a continuous positive relationship between glucose and a variety of pregnancy outcomes. Moreover, two methodologically strong studies met the inclusion criteria for this question but could not be pooled with the other studies because they examined glucose thresholds as a continuous outcome.[3,91] These studies demonstrated a continuous positive relationship between increasing glucose levels and the incidence of primary cesarean section, and macrosomia. One of these studies also found significantly fewer cases of preeclampsia, cesarean section, shoulder dystocia and/or birth injury, clinical neonatal hypoglycemia, and hyperbilirubinemia for women with no GDM compared with those meeting IADPSG criteria.[3] The clinical significance of absolute differences in event rates requires contemplation by decision makers even though statistical significance was reached at the strictest diagnostic glucose thresholds for some outcomes.

This question focused on outcomes for women who did not receive treatment for GDM. While women with untreated GDM have a variety of poorer outcomes than women without GDM, it cannot be assumed that treatment of GDM reverses all the short- and long-term poor outcomes observed in women with untreated GDM. Some of the reasons for the poorer outcomes in women that have untreated GDM may not be modifiable, such as the influences of genetic makeup. The strength of evidence was insufficient for most outcomes and comparisons in this question due to high risk of bias (observational studies), inconsistency across studies, and/or imprecise results.

Table 22. Summary of strength of evidence for the association between different glucose levels and neonatal/infant outcomes (Key Question 3)

Outcome		Number of Studies	Strength of Evidence	Summary
Macrosomia >4,000 g	CC GDM vs. no GDM	10 cohorts	Low	Statistically significant difference with fewer cases in the patient group with no GDM (RR 1.61, 95% CI 1.35, 1.92)
	CC GDM vs. false positive	5 cohorts	Low	Statistically significant difference with fewer cases in the false-positive group (RR 1.36, 95% CI 1.10, 1.68)
	CC GDM vs. 1 abnormal OGTT	3 cohorts	Low	No statistically significant difference (RR 0.99, 95% CI 0.92, 1.07)
	CC 1 abnormal OGTT vs. no GDM	7 cohorts	Low	Statistically significant difference with fewer cases in the patient group with no GDM (RR 1.44, 95% CI 1.13, 1.82)
	CC false positive vs. no GDM	5 cohorts	Low	No statistically significant difference (RR 1.02, 95% CI 0.85, 1.24)
	CC 1 abnormal OGTT vs. false positive	3 cohorts	Insufficient	-
	NDDG GDM (unrecognized) vs. no GDM	1 cohort	Insufficient	-
	NDDG false positive vs. no GDM	4 cohorts	Low	Statistically significant difference with fewer cases in the patient group with no GDM (RR 1.44, 95% CI 1.10, 1.89)
	WHO GDM vs. no GDM	1 cohort	Insufficient	-
	WHO IGT vs. no GDM	1 cohort	Insufficient	-
	IADPSG GDM vs. no GDM	2 cohorts	Insufficient	-
Macrosomia >4,500 g	CC GDM vs. no GDM	3 cohorts	Low	Statistically significant difference with fewer cases in the patient group with no GDM (RR 2.52, 95% CI 1.65, 3.84)
	CC GDM vs. false positive	2 cohorts	Insufficient	-
	CC false positive vs. no GDM	2 cohorts	Insufficient	-
	NDDG GDM (unrecognized) vs. no GDM	1 cohort	Insufficient	-

Table 22. Summary of strength of evidence for the association between different glucose levels and neonatal/infant outcomes (Key Question 3) (continued)

Outcome	Comparison	Number of Studies	Strength of Evidence	Summary
Shoulder Dystocia	CC GDM vs. no GDM	5 cohorts	Low	Statistically significant difference with fewer cases in the patient group with no GDM (RR 2.86, 95% CI 1.81, 4.51)
	CC GDM vs. false positive	1 cohort	Insufficient	-
	CC 1 abnormal OGTT vs. no GDM	1 cohort	Insufficient	-
	CC 1 abnormal OGTT vs. false positive	1 cohort	Insufficient	-
	NDDG GDM (unrecognized) vs. no GDM	1 cohort	Insufficient	-
	NDDG false positive vs. no GDM	1 cohort	Insufficient	-
	WHO IGT vs. no GDM	1 cohort	Insufficient	-
	IADPSG IGT vs. no GDM	1 cohort	Insufficient	-
	IADPSG IFG vs. no GDM	1 cohort	Insufficient	-
	IADPSG IGT-2 vs. no GDM	1 cohort	Insufficient	-
	IADPSG IGT IFG vs. no GDM	1 cohort	Insufficient	-
	IADPSG IGT vs. IFG	1 cohort	Insufficient	-
	IADPSG IGT vs. IGT-2	1 cohort	Insufficient	-
	IADPSG IGT vs. IGT IFG	1 cohort	Insufficient	-
	IADPSG IFG vs. IGT-2	1 cohort	Insufficient	-
	IADPSG IFG vs. IGT IFG	1 cohort	Insufficient	-
	IADPSG IGT-2 vs. IGT IFG	1 cohort	Insufficient	-
Neonatal Hypoglycemia	CC GDM vs. no GDM	3 cohorts	Insufficient	-
	CC GDM vs. 1 abnormal OGTT	1 cohort	Insufficient	-
	CC 1 abnormal OGTT vs. no GDM	4 cohorts	Insufficient	-
	NDDG GDM vs. no GDM	1 cohort	Insufficient	-
	NDDG false positive vs. no GDM	1 cohort	Insufficient	-
	WHO IGT vs. no GDM	3 cohorts	Insufficient	-

CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; IFG = impaired fasting glucose; IGT = impaired glucose tolerance; IGT-2 = double impaired glucose tolerance; IADPSG = International Association of Diabetes and Pregnancy Study Groups; NDDG = National Diabetes Data Group; OGTT = oral glucose tolerance test; RR = risk ratio; WHO = World Health Organization

Key Question 4

Eleven studies provided data for Key Question 4 to assess the impact of treatment for GDM on health outcomes of mothers and offspring. All studies compared diet modification, glucose monitoring, and insulin as needed with standard care. The strength of evidence for key outcomes is summarized in Table 23.

There was moderate evidence showing a significant difference for preeclampsia with fewer cases in the treated group. There was inconsistency across studies in terms of differences in maternal weight gain and the strength of evidence was considered insufficient. There were no data on long-term outcomes among women including type 2 diabetes mellitus, obesity, and hypertension.

In terms of infant outcomes, there was insufficient evidence to make a conclusion for birth injury. This was driven by lack of precision in the effect estimates and inconsistency across studies: there was no difference for RCTs but a significant difference favoring treatment in the one cohort study. The incidence of shoulder dystocia was significantly lower in the treated groups, and this finding was consistent for the 3 RCTs and 4 cohort studies. Overall, the evidence for shoulder dystocia was considered moderate showing a difference in favor of the treated group. For neonatal hypoglycemia, the strength of evidence was low suggesting no difference between groups. There was moderate evidence showing significantly lower incidence of macrosomia among the treated groups.

Only one study provided data on long-term metabolic outcomes among the offspring at 7 to 11 year followup. The strength of evidence was insufficient to reach a conclusion. For both outcomes—impaired glucose tolerance and type 2 diabetes mellitus—no differences were found between groups although the estimates were imprecise. No differences were observed in single studies that assessed BMI >95 (7-11 year followup) and BMI >85 percentile (5-7 year followup). Overall, pooled results showed no difference in offspring BMI and the strength of evidence was considered low.

In summary, there was moderate evidence showing differences in preeclampsia and shoulder dystocia with fewer cases among women (and offspring) who were treated compared with those not receiving treatment. There was also moderate evidence showing significantly fewer cases of macrosomia (>4,000 g) among offspring of women who received treatment for GDM. The results were driven by the two largest RCTs, the Maternal Fetal Medicine Unit (MFMU)[54] and the Australian Carbohydrate Intolerance in Pregnancy Study (ACHOIS),[50] which had unclear and low risk of bias, respectively. There was little evidence showing differences in other key maternal and infant outcomes between groups. One potential explanation is that for the most part the study populations included women whose glucose intolerance was less marked, as those whose glucose intolerance was more pronounced would not have been entered into a trial where they may be assigned to a group receiving no treatment. For outcomes where results were inconsistent between studies, different study glucose threshold entry criteria did not explain the variation. For some outcomes, particularly the long-term outcomes, the strength of evidence was insufficient or low suggesting that further research may change the results and increase our confidence in the results. Moreover, for some outcomes events were rare and the studies may not have had the power to detect clinically important differences between groups; therefore, findings of no significant difference should not be interpreted as equivalence between groups.

Table 23. Summary of strength of evidence for benefits of treatment (Key Question 4)

	Outcome	Number of Studies	Strength of Evidence	Summary
Maternal outcomes	Preeclampsia	3 RCTs, 1 cohort	Moderate	Significant difference in favor of treatment for RCTs (RR 0.62, 95% CI 0.43, 0.89). No difference observed for cohort study.
	Maternal weight gain	4 RCTs, 2 cohorts	Insufficient	Results not pooled for RCTs due to substantial heterogeneity. No difference for cohort studies (MD -1.04, 95% CI -2.89, 0.81).
Infant outcomes	Birth injury	2 RCTs, 1 cohort	Insufficient	No difference for RCTs (RR 0.48, 95% CI 0.12, 1.90). Significant difference favoring treatment for cohort study (RR 0.02, 95% CI 0.00, 0.22).
	Shoulder dystocia	3 RCTs, 4 cohorts	Moderate	Significant difference in favor of treatment for RCTs (RR 0.42, 95% CI 0.23, 0.77) and cohort studies (RR 0.38, 95% CI 0.19, 0.78).
	Neonatal hypoglycemia	4 RCTs, 2 cohorts	Low	No difference for RCTs (RR 1.18, 95% CI 0.92, 1.52) or cohort studies (RR 0.55, 95% CI 0.10, 2.97).
	Macrosomia (>4,000 g)	5 RCTs, 6 cohorts	Moderate	Significant difference in favor of treatment for RCTs (RR 0.50, 95% CI 0.35, 0.71). Results not pooled for cohort studies due to substantial heterogeneity.
Long-term metabolic outcomes in offspring	Impaired glucose tolerance	1 RCT	Insufficient	No difference between groups (RR 5.63, 95% CI 0.31, 101.32).
	Type 2 diabetes mellitus	1 RCT	Insufficient	No difference between groups (RR 1.88, 95% CI 0.08, 44.76).
	BMI	2 RCTs	Low	No difference between groups (RR 1.26, 95% CI 0.86, 1.84)

BMI = body mass index; CI = confidence interval; MD = mean difference; RCT = randomized controlled trial; RR = risk ratio

Key Question 5

Five studies provided data for Key Question 5 on the harms associated with treatment of GDM. There was no evidence for some of the outcomes stipulated in the protocol including costs and resource allocation.

Four of the studies provided data on the incidence of infants that were small for gestational age and showed no significant difference between groups. This finding may have resulted from inadequate power to detect differences due to a small number of events; therefore, the finding of no significant difference should not be interpreted as equivalence between groups. Four studies provided data on admission to the neonatal intensive care unit (NICU) and showed no significant differences overall. One study was an outlier as it showed significantly fewer NICU admissions in the group receiving no treatment. This difference may be attributable to site-specific policies and procedures. Two studies reported on the number of prenatal visits and generally found significantly more visits among the treatment groups.

Two RCTs showed no significant difference overall in the rate of induction of labor, although there was important statistical heterogeneity between studies. One RCT showed significantly more inductions of labor in the treatment group[50] while the other study did not.[54] Different study protocols may account for the heterogeneity of results between studies. In the

first study, that showed more inductions of labor in the treatment group, no recommendations were provided regarding obstetrical care. In the later study, antenatal surveillance was reserved for standard obstetrical indications. Based on the studies included in Key Question 4, there was no difference in Cesarean section between treatment and non treatment GDM (5 RCTs and 6 cohort studies).

A single study assessed depression and anxiety at 6 weeks after study entry and 3 months postpartum using the Spielberger State-Trait Anxiety Inventory and the Edinburgh Postnatal Depression Score, respectively. There was no significant difference between groups in anxiety at either time point, although there were significantly lower rates of depression in the treatment group at 3 months postpartum. These results should be interpreted cautiously because the assessment of depression and anxiety was conducted in a subgroup of the larger RCT. Maternal stress in pregnancy has been associated with poor metabolic consequences in offspring.[168] Other research found that women with GDM compared with glucose tolerant women had a higher level of anxiety at time of the first assessment; however, before delivery these differences in anxiety scores did not persist.[169]

Findings in Relationship to What Is Already Known

This review provides evidence that treating GDM reduces some poor maternal and neonatal outcomes. The recent randomized trial published in 2009 by the MFMU[54] reinforces the findings of the earlier ACHOIS trial which was published in 2005 [50] and included in an earlier version of this review.[53] Both trials showed that treating GDM to targets of 5.3 or 5.5 mmol/L fasting and 6.7 or 7.0 mmol/L 2 hours post-meal reduced neonatal birthweight, large for gestational age, macrosomia, shoulder dystocia, and preeclampsia without a reduction in neonatal hypoglycemia or hyperbilirubinemia/jaundice requiring phototherapy, or an increase in small for gestational age. In contrast to the ACHOIS trial, MFMU demonstrated a reduced cesarean section rate in the GDM treatment group. The failure of ACHOIS to find a lower cesarean section rate despite reduced neonatal birthweight and macrosomia may have been the result of differing obstetrical practices or the different populations studied (e.g., the inclusion of some women with more marked glucose intolerance in ACHOIS as reflected by the increased prevalence of insulin use; more black and Hispanic women in the MFMU study). Differences may have also resulted due to study design: in ACHOIS, participants did not receive specific recommendations regarding obstetrical care, thus replicating obstetrical care for women with GDM. In the MFMU study, antenatal surveillance was reserved for standard obstetrical indications. Our findings of the effect of treatment of GDM is similar to a systematic review and meta-analysis published in 2010 by Horvath et al.[170] that included two older RCTs of GDM that were not included in our analysis because we restricted our inclusion criteria to studies published after 1995.

The Hyperglycemia and Adverse Pregnancy Outcomes (HAPO) Study Cooperative Research Group [3] confirmed findings of the earlier Toronto Trihospital study [142] in a large international sample of women with a simpler 75 g OGTT showing a continuous positive association between maternal glucose and increased birthweight, as well as fetal hyperinsulinemia (HAPO only), at levels below diagnostic thresholds for GDM that existed at the time of the study. However, no clear glucose thresholds were found for fetal overgrowth or a variety of other maternal and neonatal outcomes. Subsequently, the IADPSG developed diagnostic thresholds for GDM based on a consensus of expert opinion of what was considered to be the most important outcomes and the degree of acceptable risk for these outcomes. The thresholds chosen by the IADPSG were derived from the HAPO data to identify women with a higher risk (adjusted odds ratio 1.75) of

large for gestational age, elevated c-peptide, high neonatal body fat compared with the mean maternal glucose values of the HAPO study. The glucose threshold chosen by the IADPSG represents differing levels of risk for other outcomes. Specifically the IADPSG thresholds represent a 1.4 (1.26-1.56) risk for pregnancy induced hypertension and a 1.3 (1.07,1.58) risk for shoulder dystocia.

Neither recent RCT was designed to determine diagnostic thresholds for GDM or therapeutic glucose targets. However, it is noteworthy that therapeutic glucose targets for both ACHOIS and MFMU were above the proposed diagnostic criteria of the IADPSG (fasting 5.5 mmol/L (99 mg/dL) and 5.3 mmol/L (95 mg/dL and 2 hour post-meal of 7.0 mmol/L (126 mg/dL and 6.7 mmol/L 120 mg/dL), respectively). A change in diagnostic criteria without addressing management thresholds could contribute to clinical confusion. If diagnostic thresholds for GDM below the treatment targets of the large RCTs are endorsed, this could ethically obstruct the possibility of future RCTs to compare different treatment targets above such diagnostic thresholds.

It has been hypothesized that treatment of GDM may reduce future poor metabolic outcomes for children born to mothers with GDM. If true, the potential for long-term gain is important from a clinical and public health perspective and may justify the "costs" of screening and treating women for GDM. However, the followup of offspring from two RCTs [50,96] and a HAPO cohort in Belfast [171] currently fail to support this hypothesis. This may be explained in part due to insufficient length of followup or inadequate numbers of events.

The HAPO study showed that maternal weight and glucose predict large for gestational age. However, body mass index was the better predictor of large for gestational age than glucose until glucose thresholds higher than the diagnostic thresholds set by the IADPSG were reached.[172,173] Most cases of large for gestational age occur in neonates of mothers with normal glycemia. A large observational study found that the upper quartile of maternal BMI accounted for 23 percent of macrosomia, while GDM was responsible for only 3.8 percent.[174] The ongoing obesity epidemic in the United States warrants careful consideration of a diagnostic approach for GDM that incorporates maternal BMI. This would require the development and validation of a risk model that incorporates maternal BMI as well as other modifiable risk factors. Such a model could facilitate the identification of women at high risk of adverse pregnancy outcomes and minimize exposure of lower risk women to unnecessary interventions.

Applicability

There are several issues that may limit the applicability of the evidence presented in this review to the U.S. population, and these vary slightly by Key Question. All of the Key Questions asked about the effects of screening and treatment before and after 24 weeks' gestation. The vast majority of included studies screened women after 24 weeks' gestation, therefore the results are not applicable to screening and treatment earlier in gestation.

For Key Question 1 on the test properties of screening and diagnostic tests, comparisons involving the WHO criteria are less applicable to the U.S. setting as these criteria are not used in North America. There were insufficient data from the included studies to assess the performance of screening or diagnostic tests for specific patient characteristics (e.g., BMI, race/ethnicity). Therefore it is unclear whether the evidence applies to specific subpopulations of women.

For Key Question 2, limited evidence was identified because the comparison of interest was women who had not undergone screening. As screening is routine in prenatal care in the United States, the evidence (or limited evidence) is likely not helpful for U.S. decisionmaking and a

refinement of this question may be appropriate to reflect current practices and outstanding questions.

With respect to Key Question 3, all studies or groups included for analysis involved women who had not received treatment for GDM. It cannot be assumed that the same association and outcomes would be observed in clinical practice where standard care is to screen for and treat GDM. The untreated women may differ from the general population in ways that are related to the reasons for which they did not seek or receive early prenatal care (e.g., socioeconomic status). That is, the reasons that they did not receive treatment for GDM are varied; some reasons such as late presentation for obstetrical care may confound the observed association with health outcomes. Attempts were made to control for these factors in some studies by including a group of women without GDM with similar known confounders[e.g.,146] or by adjusting for known confounders in the analysis. The adjusted estimates did not change the overall pooled results in the majority of cases and did not change the overall conclusions.

The majority of the studies for Key Questions 4 and 5 pertaining to the benefits and harms of treatment for GDM were conducted in North America or Australia. Most of the North American studies were inclusive of mixed racial populations and are likely applicable to the general U.S. population. Even though the Australian RCT[50] population had more white women with a lower BMI than the U.S. RCT (MFMU[54]), this should not affect applicability of most of their findings because these patient characteristics would be factors associated with lower risk of poor outcomes. Differences in physician or hospital billing structures between the United States and Australia may have accounted for the discrepant findings with respect to NICU admissions and as a result limit the applicability of this finding in the United States. Among the studies included in Key Questions 4 and 5, a variety of glucose threshold criteria were used for inclusion, varying from 50 g screen positive with nondiagnostic oral glucose tolerance tests to women who met National Diabetes Data Group criteria for a diagnosis of GDM. The two large RCTs used different glucose thresholds for entry in their trials: WHO and CC criteria with a fasting glucose <95 mg/dL (5.3 mmol/L), respectively.[50,54] The mean glucose levels at study entry were similar between these two RCTs, which may reflect a reluctance to assign women with more marked glucose intolerance to a group receiving no treatment. The results may not be applicable to women with higher levels of glucose intolerance.

Limitations of the Evidence Base

There is sparse evidence to clarify issues regarding the timing of screening and treatment for GDM (i.e., before and after 24 weeks' gestation). Earlier screening will help identify overt type 2 diabetes mellitus and distinguish this from GDM. This has important implications for clinical management and ongoing followup beyond pregnancy. Previously unrecognized type 2 diabetes mellitus diagnosed in pregnancy should be excluded from the diagnosis of GDM because this condition has the highest perinatal mortality rate of all classes of glucose intolerance in pregnancy.[175] This distinction within research studies will provide more targeted evidence to assist obstetrical care providers to risk stratify obstetrical care and glycemic management of patients with overt type 2 diabetes mellitus diagnosed in pregnancy and those with less pronounced pregnancy-induced glucose intolerance. This will also facilitate better comparability across future studies. There were few data available on long-term outcomes. Furthermore, the studies included in this review do not provide evidence of a direct link between short-term and long-term outcomes (e.g., macrosomia and childhood obesity).

Care provider knowledge of the glucose screening and diagnostic results may have introduced a bias if their subsequent treatment of women differed depending on the results. This was of particular concern for Key Question 3. For Key Question 3, which assessed how the various criteria for GDM influenced pregnancy outcomes, many of the statistically significant differences seemed to be driven by the size of the study or pooled analysis, i.e., statistically significant differences could be found if the sample were sufficiently large. However, these differences may not be clinically important. The absolute differences in event rates between different glucose thresholds need careful consideration by decisionmakers even though statistically significant differences were found. Another key limitation with the evidence for Key Question 3 is that the studies included were cohort studies, many of which did not control for potential confounders. Therefore, any associations between glucose thresholds and outcomes should be interpreted with caution.

Given that the large landmark studies[91,142] show a continuous relationship between glucose and maternal and neonatal outcomes, the lack of clear thresholds contributes to the uncertainty regarding a diagnostic threshold for GDM. While there is controversy about where to set lower limits for diagnostic criteria, the identification of overt diabetes in pregnancy is imperative if this diagnosis has not occurred prior to pregnancy. Overt diabetes first identified in pregnancy should be distinguished from GDM in order to gain a better understanding of the true risk of GDM to pregnancy outcomes. Unfortunately there is no literature to guide diagnostic criteria for a diagnosis of overt diabetes in pregnancy.

There were several methodological concerns for this evidence base. For example, risk of spectrum bias and partial verification bias (Key Question 1); different definitions or methods of assessing key outcomes (e.g., clinical vs. biochemical neonatal hypoglycemia and hyperbilirubinemia) (Key Questions 3 and 4); and, lack of blinding of treatment arms in some studies (Key Questions 4 and 5).

Future Research

Several important gaps in the current literature exist:

- The adoption of a consistent comparator for diagnosis of GDM, such as the 75 g OGTT, would facilitate comparisons across studies even if different diagnostic thresholds are used.
- Further analysis of the HAPO data could help answer some outstanding questions. For example, further analysis could better define absolute differences in rare event rates. This evidence could be used to inform discussions about the clinical importance of absolute differences in event rates at thresholds other than those of the IADPSG. Such analyses should include adjustment for important confounders such as maternal BMI.
- Further analysis of the HAPO data examining center to center differences in glucose outcome relationships would be helpful in determining the usefulness of FPG as a screening test for GDM.
- Research is needed to clarify issues regarding earlier screening and treatment, particularly as they relate to the diagnosis, treatment, and long-term outcomes of pregestational (overt) diabetes.
- FPG is a screening test that requires further research, given that the reproducibility of fasting glucose measurement is superior to post glucose load measurements.[165]
- Further study of the long-term metabolic impact on offspring whose mothers have been treated for GDM is warranted. In addition, data on the influences of GDM treatment on

long-term breastfeeding success have not been studied. The association of breastfeeding with reduced poor metabolic outcomes in offspring of GDM has been found to have a dose dependent response with duration of breastfeeding.[176]

- Well-conducted prospective cohort studies of the "real world" impact of GDM treatment on care utilization are needed.
- Research is needed to help determine the glucose thresholds and treatment targets at which GDM treatment benefits outweigh the risks of treatment and no treatment. This will best be achieved through well-conducted, large RCTs that randomize women with GDM to different glucose treatment targets.
- While this review did not identify evidence of substantial harms to treatment, the populations considered were mostly women whose GDM was controlled without medication. There is a risk for more precautionary management of women diagnosed with GDM who are perceived by clinicians to be at greater risk, such as those managed with insulin, which may result in unnecessary interventions (e.g., cesarean section).[177] Therefore, RCTs investigating the care of women diagnosed with GDM, including fetal surveillance protocols, are needed to guide obstetrical investigations and management of GDM. Further, RCTs comparing delivery management for GDM with and without insulin or medical management are needed to provide clinicians guidance on appropriate timing and management of delivery in women with GDM to avoid unnecessary intervention in "the real world" driven by health care provider apprehension.
- Long-term studies that evaluate the potential increased or decreased resource utilization associated with the implementation of diabetes prevention strategies after a diagnosis of GDM are required.
- Studies to assess the long-term impact that a label of GDM may have for future pregnancy planning, future pregnancy management, and future insurability are required.
- The increased prevalence of type 2 diabetes mellitus in women of reproductive age merits consideration of preconception screening for overt diabetes in women at risk of type 2 diabetes. In addition to poor maternal and neonatal outcomes associated with overt diabetes in pregnancy, there is potential for benefit of preconception care.
- Long-term benefits and harms need to be evaluated among different treatment modalities for GDM (e.g., diet, exercise, insulin, oral glucose lowering medications, and/or combinations of these).
- Since 2011-2012 the ADA has endorsed the use of an HbA1c of 6.5 percent or more as diagnostic of diabetes in nonpregnant women.[36] Studies of HbA1c with trimester-specific cutoffs to determine the value at which overt diabetes should be diagnosed in pregnancy are needed.

Limitations of the Review

This review followed rigorous methodological standards which were detailed a priori. The limitations of the review to fully answer the Key Questions are largely due to the nature and limitations of the existing evidence.

There are several limitations that need to be discussed regarding systematic reviews in general. First, there is a possibility of publication bias. The impact of publication bias on the results of diagnostic test accuracy reviews (Key Question 1) is not well understood nor have the tools to investigate publication bias in these reviews been developed. For the remaining Key Questions we may be missing unpublished and/or negative therapy studies, and may be

overestimating the benefits of certain approaches. However, we conducted a comprehensive and systematic search of the published literature for potentially relevant studies. Search strategies included combinations of subject headings and free text words. These searches were supplemented by handsearching for gray literature (i.e., unpublished or difficult to find studies). Despite these efforts, we recognize that we may have missed some studies.

There is also a possibility of study selection bias. However, we employed at least two independent reviewers and feel confident that the studies that were excluded from this report were done so for consistent and appropriate reasons. Our search was comprehensive, so it is unlikely that there are many studies in press or publication that were missed.

Cost analysis of different screening and diagnostic approaches was not addressed in this review.

Conclusions

There was limited evidence regarding the test characteristics of current screening and diagnostic strategies for GDM. Lack of an agreed upon gold standard for diagnosis of GDM creates challenges for assessing the accuracy of tests and comparing across studies. The 50 g OGCT with a glucose threshold of 130 mg/dL versus 140 mg/dL improves sensitivity and reduces specificity (10 studies). Both thresholds have high NPV, but variable PPV across a range of GDM prevalence. There was limited evidence for the screening of GDM diagnosed less than 24 weeks' gestation (3 studies). Single studies compared the diagnostic characteristics of different pairs of diagnostic criteria in the same population. The use of fasting glucose (≥85 mg/dL) as a screen for GDM may be a practical alternative because of similar test characteristics to the OGCT particularly in women who cannot tolerate any form of oral glucose load.

Evidence supports benefits of treating GDM with little evidence of short-term harm. Specifically, treatment of GDM results in lower incidence of preeclampsia, macrosomia, and large for gestational age infants. Current research does not demonstrate a treatment effect of GDM on clinical neonatal hypoglycemia or future poor metabolic outcomes of the offspring. RCTs of GDM treatment show limited harm related to treating GDM, other than an increased demand for services. There is a risk for more precautionary management of women diagnosed with GDM who are perceived by clinicians to be at greater risk, such as those managed with insulin, which may result in unnecessary interventions (e.g., cesarean section); however, this review found limited data for these outcomes and further research on the care of women diagnosed with GDM (e.g., fetal surveillance protocols) is warranted.

What remains less clear is what the lower limit diagnostic thresholds for GDM should be. Given the continuous association between glucose and a variety of outcomes, decisions should be made in light of what outcomes that are altered by treatment are most important and what level of increased risk is acceptable. A dichotomous view of GDM may no longer be appropriate, given evidence of a continuous relationship between maternal blood glucose and pregnancy outcomes. An alternative approach would be to define different glucose thresholds based on maternal risk for poor pregnancy outcomes.

Further study is needed regarding the long-term metabolic impact on offspring of mothers receiving GDM treatment; the "real world" impact of GDM treatment on care utilization outside of structured research trials; and, the impact of the timing of screening for GDM, particularly before 24 weeks' gestation and in the first trimester of pregnancy. Early screening could help identify pregestational (i.e., overt) diabetes. Research is urgently required to determine the best

way to diagnose and manage overt diabetes in pregnancy, particularly in an era of increasing rates of obesity and diabetes in the U.S. population.

Table 24. Summary of Evidence for all Key Questions

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ1. What are the sensitivities, specificities, reliabilities, and yields of current screening tests for GDM? (A) After 24 weeks' gestation? (B) During the first trimester and up to 24 weeks' gestation?	***A) After 24 wk gestation*** 51 prospective studies *Fair to good quality*	Limitations: Lack of an agreed upon gold standard for diagnosis of GDM creates challenges for assessing the accuracy of tests and comparing across studies. GDM was confirmed using criteria developed by CC, ADA, NDDG, and WHO. There were sparse data comparing overall approaches for diagnosis and screening, e.g., one-step vs. two-step, selective vs. universal. Consistency: Across studies, numerous tests and thresholds were examined. Screening tests included the 50 g OGCT, FPG risk factor-based screening, and other less common tests such as HbA1c, serum fructosamine.	Prevalence of GDM varied across studies and diagnostic criteria used. Results need to be interpreted in the context of prevalence. Comparisons involving WHO criteria are less applicable to the North American setting because these criteria are not used in North America.	• Prevalence varied across studies and diagnostic criteria: ADA 2000-2010 (75 g) 2.0 to 19% (range), CC 3.6 to 38%, NDDG 1.4 to 50%, WHO 2 to 24.5%. • 9 studies examined a 50 g OGCT with a cutoff value of ≥140 mg/dL; GDM was confirmed using CC criteria. Results: sensitivity 85%, specificity 86%, prevalence 3.8 to 31.9%, PPV 18 to 27% (prevalence <10), PPV 32 to 83% (prevalence ≥10), NPV median 98%. • 6 studies examined a 50 g OGCT (≥130 mg/dL); GDM was confirmed using CC criteria. Results: sensitivity 99%, specificity 77%, prevalence 4.3 to 29.5%, PPV 11 to 31% (prevalence <10), PPV 31 to 62% (prevalence ≥10), NPV median 100%. • 1 study examined a 50 g OGCT (≥200 mg/dL); GDM was confirmed using CC criteria. Sensitivity, specificity, PPV, and NPV were all 100%. Prevalence was 6.4%. • 7 studies examined a 50 g OGCT (≥140 mg/dL); GDM was confirmed using NDDG criteria. Results: sensitivity 85%, specificity 83%, prevalence 1.4 to 45.8%, PPV 12 to 39% (prevalence <10), PPV 57% (prevalence ≥10), NPV median 99%. • 3 studies examined a 50 g OGCT (≥130 mg/dL); GDM was confirmed using NDDG criteria. Results: sensitivity 67 to 90% (range), specificity 47 to 84%, prevalence 16.7 to 35.3%, PPV 20 to 75%, NPV 86 to 95%. • 3 studies examined a 50 g OGCT (different thresholds); GDM was confirmed using ADA 2000-2010 (75 g) criteria. Prevalence was 1.6 to 4.1 (range). Results: sensitivity 86 to 97% (range), specificity 79 to 87%, PPV 7 to 20%, NPV 99 to 100%.

Table 24. Summary of Evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ1. What are the sensitivities, specificities, reliabilities, and yields of current screening tests for GDM? (A) After 24 weeks' gestation? (B) During the first trimester and up to 24 weeks' gestation? (continued)	***A) After 24 wk gestation*** 51 prospective studies *Fair to good quality* (continued)			• 3 studies examined a 50 g OGCT (≥140 mg/dL); GDM was confirmed using WHO criteria. Results: sensitivity 43 to 85%, specificity 73 to 94%, prevalence 3.7 to 15.7%, PPV 18 to 20% (prevalence <10), PPV 58% (prevalence ≥10), NPV median 99%. • 7 studies examined FPG at different thresholds; GDM was confirmed using CC criteria. Results: at ≥85 mg/dL sensitivity 87%, specificity 52%; at ≥90 mg/dL sensitivity 77%, specificity 76%; at ≥92 mg/dL sensitivity 76%, specificity 92%; at ≥95 mg/dL sensitivity 54%, specificity 93%. At ≥85 mg/dL prevalence 1.4 to 34.53 (range). PPV 10% (prevalence <10) and 23 to 59% (prevalence ≥10). Median NPV 93%. • 8 studies examined risk factor-based screening but were not pooled. Studies used different criteria to confirm GDM. Results: sensitivity 48 to 95% (range), specificity 22 to 94%, prevalence 1.7 to 16.9%, PPV 5 to 19% (prevalence <10), PPV 20% (prevalence ≥10), NPV median 99%. • 1 study compared IADPSG vs. ADIPS 2 step (reference) to diagnose GDM. Results: sensitivity 82%, specificity 94%, prevalence 13.0%, PPV 61%, NPV 98%. • 4 studies compared 75 g and 100 g load tests to diagnose GDM. Prevalence ranged from 1.4 to 50%. Results were not pooled: sensitivity 18 to 100%, specificity 86 to 100%, PPV 12 to 100%, NPV 62 to 100%.

Table 24. Summary of Evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ1. What are the sensitivities, specificities, reliabilities, and yields of current screening tests for GDM? (A) After 24 weeks' gestation? (B) During the first trimester and up to 24 weeks' gestation? (continued)	***(B) During the first trimester and up to 24 wk gestation*** 3 prospective cohort studies	Limitations: Only 3 studies of women before 24 wks gestation; therefore, no conclusions can be made for test characteristics in early pregnancy. Consistency: Not applicable (not enough studies addressing the same question to judge consistency).	Evidence too limited to judge applicability.	• 1 study examined the 50 g OGCT at 10 wks and confirmed GDM using JSOG criteria (75 g). Results: sensitivity 88%, specificity 79%, prevalence 1.6%, PPV 7%, NPV 100%. • 1 study examined 50 g OGCT at 20 wks and confirmed GDM using ADA (2000-2010) 100 g criteria. Results: sensitivity 56%, specificity 94%, prevalence 3.6%, PPV 24%, NPV 98%. • 1 study compared 1st and 2nd trimester results using 3 screening tests (OGCT at ≥130 mg/dL, FPG, HbA1c); GDM confirmed using JSOG criteria. Results (OGCT) 1st trimester: prevalence 1.9%, sensitivity 93%, specificity 77%, PPV 7.1, NPV 99%; 2nd trimester: prevalence 2.9%, sensitivity 100%, specificity 85%, PPV 17%, NPV 100%.
KQ2: What is the direct evidence on the benefits and harms of screening women (before and after 24 weeks' gestation) for GDM to reduce maternal, fetal, and infant morbidity and mortality?	2 retrospective cohort studies *Fair and good quality*	Limitations: No RCTs available to answer this question. Consistency: Not applicable (not enough studies addressing the same question to judge consistency).	The comparison for this question was women who had and had not undergone screening. Since screening is now commonplace it may be unlikely to identify studies or cohorts where this comparison is feasible.	• 1 study (n=1,000) showed more cesarean deliveries in the screened group. A second study (n=93) found the incidence of macrosomia (≥4.3 kg) was the same in screened and unscreened groups (7% each group). • Based on the small number of studies and sample sizes, the effect of screening women for GDM on health outcomes is inconclusive.

Table 24. Summary of Evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ3: In the absence of treatment, how do health outcomes of mothers who meet various criteria for GDM and their offspring compare to those who do not meet the various criteria?	38 prospective or retrospective cohort studies; 2 studies were long-term followup from RCTs; however, only data from the untreated patients were included. *Fair to good quality*	Limitations: Strength of evidence was low to insufficient for all graded outcomes due to risk of bias (all observational studies), inconsistency, and/or imprecision. For many comparisons, the numbers of studies, participants, and/or events was low; therefore, findings of no statistically significant differences between groups do not imply equivalence or rule out potential differences. Consistency: A wide variety of diagnostic criteria and thresholds were compared across studies. There were often few studies with similar comparison groups. Differences in defining and assessing outcomes may have contributed to heterogeneity in results across studies (e.g., biochemical vs. clinical assessment of neonatal hypoglycemia).	All studies or groups included for analysis involved women who had not received treatment for GDM. These women may differ from the general population in other ways that are related to the reasons that they did not seek or receive early prenatal care (e.g., socioeconomic status).	*Maternal outcomes:* • A methodologically strong study showed a continuous positive relationship between increasing glucose levels and the incidence of primary cesarean section. This study also found significantly fewer cases of preeclampsia and cesarean section for women with no GDM vs. IADPSG. • For preeclampsia, significant differences were found for CC vs. patients with no GDM (3 studies), with fewer cases among the patients with no GDM, and for CC vs. false-positive groups (2 studies), with fewer cases among the false positives. The strength of evidence was low. No differences were found for NDDG false positive (2 studies), NDDG 1 abnormal OGTT vs. no GDM (1 study), or IGT WHO vs. no GDM (3 studies); the strength of evidence was insufficient. • For maternal weight gain, significant differences were found for 3 of 12 comparisons: IADPSG IGT vs. no GDM (favored IGT), IADPSG IFG vs. no GDM (favored IFG), IADPSG IGT-2 vs. no GDM (favored IGT-2). All comparisons were based on single studies (strength of evidence insufficient). *Fetal/neonatal/child outcomes:* • 2 methodologically strong studies showed a continuous positive relationship between increasing glucose levels and the incidence of macrosomia. 1 of these studies also showed significantly fewer cases of shoulder dystocia and/or birth injury, clinical neonatal hypoglycemia, and hyperbilirubinemia for women with no GDM vs. IADPSG.

Table 24. Summary of Evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ3: In the absence of treatment, how do health outcomes of mothers who meet various criteria for GDM and their offspring compare to those who do not meet the various criteria? (continued)				• For macrosomia >4,000 g, 6 of 11 comparisons showed a significant difference: patient groups with no GDM had fewer cases compared with CC GDM (10 studies), CC 1 abnormal OGTT (7 studies), NDDG GDM (unrecognized) (1 study), NDDG false positives (4 studies), and WHO IGT (1 study). Fewer cases were found for women with false-positive results compared with CC GDM (5 studies). Data for macrosomia >4,500 g were available for 4 comparisons and showed significant differences in 2 cases: patient groups with no GDM had fewer cases compared with CC GDM (3 studies) and unrecognized NDDG GDM (1 study). The strength of evidence for macrosomia was low to insufficient. • For shoulder dystocia, significant differences were found for 7 of 17 comparisons; all comparisons but 1 were based on single studies (insufficient strength of evidence). Patient groups with no GDM showed lower incidence of shoulder dystocia when compared with CC GDM (5 studies, low strength of evidence), NDDG GDM (unrecognized), NDDG false positive, WHO IGT, IADPSG IFG, and IADPSG IGT IFG. The other significant difference showed lower incidence among the false-positive group compared with CC 1 abnormal OGTT.

Table 24. Summary of Evidence for all Key Questions (continued)

Key Question	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ3: In the absence of treatment, how do health outcomes of mothers who meet various criteria for GDM and their offspring compare to those who do not meet the various criteria? (continued)				• For fetal birth trauma/injury, single studies compared CC GDM and WHO IGT with no GDM and showed no differences. Two studies showed fewer cases for no GDM compared with NDDG GDM. Strength of evidence was insufficient for all comparisons. • No differences were found for neonatal hypoglycemia for any comparison, including CC GDM vs. no GDM (3 studies), CC GDM vs. 1 abnormal OGTT (1 study), CC 1 abnormal OGTT vs. no GDM (4 studies), NDDG GDM vs. no GDM (1 study), NDDG false positive vs. no GDM (1 study), and WHO IGT vs. no GDM (3 studies). Strength of evidence was insufficient for all comparisons.
KQ4: Does treatment modify the health outcomes of mothers who meet various criteria for GDM and offspring?	5 RCTs and 6 retrospective cohort studies. *Poor to good quality*	Limitations: For some outcomes, particularly the long-term outcomes, the strength of evidence was insufficient or low. Moreover, for some outcomes events were rare and the studies may not have had the power to detect clinically important differences between groups; therefore, findings of no significant difference should not be interpreted as equivalence between groups.	For the most part, study populations included women whose glucose intolerance was less marked, as those whose glucose intolerance was more pronounced would not be entered into a trial in which they may be assigned to a group receiving no treatment. The majority of studies were conducted in North America or Australia, with 2 from Italy. Most of the North American studies were inclusive of mixed racial populations and are likely applicable to the general U.S. population. Even though the Australian RCT population had more white women with a lower	*Maternal outcomes:* • Moderate evidence from 3 RCTs showed a significant difference for preeclampsia, with fewer cases in the treated group. • There was inconsistency across studies in terms of maternal weight gain (4 RCTs and 2 cohort studies); the strength of evidence was insufficient due to inconsistency and imprecision in effect estimates. *Offspring outcomes:* • There was insufficient evidence to make a conclusion for birth injury. There was inconsistency across studies with the 2 RCTs showing no difference and the 1 cohort study showing a difference in favor of the treated group. The low number of events and participants across all studies resulted in imprecise estimates. • Moderate evidence showed significantly lower incidence of shoulder dystocia in the treated groups, and this finding was consistent for the 3 RCTs and 4 cohort studies.

Table 24. Summary of Evidence for all Key Questions (continued)

Key Questions	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ4: Does treatment modify the health outcomes of mothers who meet various criteria for GDM and offspring? (continued)		Consistency: Some inconsistency occurred at 2 levels. First, there were inconsistencies for some outcomes between RCTs and observational studies which may be attributable to confounding and methods of selecting study groups (e.g.,historical control groups). Second, in some instances there were inconsistencies across studies within designs that were often attributable to the manner in which outcomes were defined or assessed (e.g., clinical vs. biochemical assessment of neonatal hypoglycemia).	BMI than the U.S. RCT; this should not affect applicability of most of their findings for the U.S. women because these subject characteristics would be factors associated with lower risk of poor outcomes.	• There was low evidence of no difference between groups for neonatal hypoglycemia based on 4 RCTs and 2 cohort studies. • For outcomes related to birthweight (including macrosomia >4,000 g, macrosomia >4,500 g, actual birthweight, and large for gestational age), differences were often observed favoring the treated groups. Strength of evidence was moderate for macrosomia >4,000 g. • 1 RCT followed patients for 7 to 11 years and found no differences for impaired glucose tolerance or type 2 DM, although the strength of evidence was considered insufficient. • No differences were observed in single studies that assessed BMI >95 (7-11 year followup) and BMI >85 percentile (5-7 year followup). Overall, pooled results showed no difference in BMI, and the strength of evidence was considered low
KQ5: What are the harms of treating GDM and do they vary by diagnostic approach?	4 RCTs and 1 retrospective cohort study. *Fair to good quality*	*Limitations:* No study evaluated costs and resource allocation. Limited evidence on harms. Limited evidence for number of prenatal visits and NICU admissions. Findings of no significant differences may be attributable to low power and should not be interpreted as equivalence. *Consistency:* Not applicable (not enough studies addressing the same question to judge).	As above for KQ4. In addition, differences in billing structures between the United States and Australia may have accounted for the discrepant findings with respect to NICU admissions between these studies and as a result limit the applicability of this finding in the United States.	• 1 RCT assessed depression and anxiety at 6 weeks after study entry and 3 months postpartum. • There was no significant difference between groups in anxiety at either time point, although there were significantly lower rates of depression in the treatment group at 3 months postpartum. • 4 RCTs reported small for gestational age and found no significant difference. 3 RCTs and 1 cohort study provided data on admission to NICU and showed no significant differences overall. One trial was an outlier because it showed a significant difference favoring the no treatment group. This difference may be attributable to site-specific policies and procedures.

Table 24. Summary of Evidence for all Key Questions (continued)

Key Questions	Number and Quality of Studies	Limitations/ Consistency	Applicability	Summary of Findings
KQ5: What are the harms of treating GDM and do they vary by diagnostic approach? (continued)				• 2 RCTs reported on the number of prenatal visits and generally found more visits among the treatment groups. • 2 RCTs reporting on induction of labor showed different results, with 1 showing a significant difference with more cases in the treatment group and the other showing no difference. • Based on studies included in KQ4, no differences between groups were found for cesarean section (5 RCTs, 6 cohorts) or unplanned cesarean section (1 RCT, 1 cohort).

ADA = American Diabetes Association; ADIPS = Australasian Diabetes in Pregnancy Society; BMI = body mass index; CC = Carpenter-Coustan; DM = diabetes mellitus; FPG = fasting plasma glucose; GDM = gestational diabetes mellitus; HbA1c = glycated hemoglobin; IADPSG = International Association of Diabetes in Pregnancy Study Groups; IFG = impaired fasting glucose; IGT = impaired glucose tolerance; IGT-2 = double impaired glucose intolerance; JSOG = Japan Society of Obstetrics and Gynecology; KQ = Key Question; NDDG = National Diabetes Data Group; NPV = negative predictive value; NICU = neonatal intensive care unit; OGCT = oral glucose challenge test; OGTT = oral glucose tolerance test; PPV = positive predictive value; RCT = randomized controlled trial; wk(s) = week(s); WHO = World Health Organization

References

1. Balsells M, Garcia-Patterson A, Gich I, et al. Maternal and fetal outcome in women with type 2 versus type 1 diabetes mellitus: a systematic review and metaanalysis. J Clin Endocrinol Metab. 2009;94(11):4284-91. PMID: 19808847.
2. National Diabetes Data Group. Diabetes in America, 2nd ed. Bethesda, MD: National Institutes of Health; 1995.
3. HAPO Study Cooperative Research Group, Metzger B, Lowe L, et al. Hyperglycemia and Adverse Pregnancy Outcomes. N Engl J Med. 2008;358(19):1991-2002. PMID: 18463375.
4. American Diabetes Association. Position statement: standards of medical care in diabetes - 2012. Diabetes Care. 2012;35(Suppl 1):S11-S63. PMID: 22187469.
5. Carpenter MW, Coustan DR. Criteria for screening tests for gestational diabetes. Am J Obstet Gynecol. 1982;144(7):768-73. PMID: 7148898.
6. Sacks DA, Hadden DR, Maresh M, et al. Frequency of gestational diabetes mellitus at collaborating centers based on IADPSG consensus panel-recommended criteria: the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study. Diabetes Care. 2012;35(3):526-8. PMID: 22355019.
7. Ferrara A. Increasing prevalence of gestational diabetes mellitus: a public health perspective. Diabetes Care. 2007;30 Suppl 2:S141-S146. PMID: 17596462.
8. Mokdad AH, Ford ES, Bowman BA, et al. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. JAMA. 2003;289(1):76-9. PMID: 12503980.
9. Gillman MW, Oakey H, Baghurst PA, et al. Effect of treatment of gestational diabetes mellitus on obesity in the next generation. Diabetes Care. 2010;33(5):964-8. PMID: 20150300.
10. Kaufmann RC, Schleyhahn FT, Huffman DG, et al. Gestational diabetes diagnostic criteria: long-term maternal follow-up. Am J Obstet Gynecol. 1995;172(2 Pt 1):621-5. PMID: 7856695.
11. Mokdad AH, Ford ES, Bowman BA, et al. Diabetes trends in the U.S.: 1990-1998. Diabetes Care. 2000;23(9):1278-83. PMID: 10977060.
12. Mokdad AH, Serdula MK, Dietz WH, et al. The spread of the obesity epidemic in the United States, 1991-1998. JAMA. 1999;282(16):1519-22. PMID: 10546690.
13. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2006;29(Suppl 1):S43-S48. PMID: 16373932.
14. Berger H, Crane J, Farine D, et al. Screening for gestational diabetes mellitus. J Obstet Gynaecol Can. 2002;24(11):894-912. PMID: 12417905.
15. Naylor CD, Sermer M, Chen E, et al. Selective screening for gestational diabetes mellitus. Toronto Trihospital Gestational Diabetes Project Investigators. N Engl J Med. 1997;337(22):1591-6. PMID: 9371855.
16. Hillier T, Vesco K, Pedula K, et al. Screening for gestational diabetes mellitus: A systematic review for the U.S. preventive services task force. Ann Intern Med. 2008;148:766-75. PMID: 18490689
17. American College of Obstetricians and Gynecologists Committee on Practice Bulletins. ACOG Practice Bulletin. Clinical management guidelines for obstetrician-gynecologists. Number 30, September 2001. Gestational Diabetes. Obstet Gynecol. 2001;98(3):525-38. PMID: 11547793.
18. Meltzer SJ, Snyder J, Penrod JR, et al. Gestational diabetes mellitus screening and diagnosis: a prospective randomised controlled trial comparing costs of one-step and two-step methods. BJOG. 2010;117(4):407-15. PMID: 20105163.
19. Gabbe S, Gregory R, Power M, et al. Management of diabetes mellitus by obstetrician-gynecologists. Obstet Gynecol. 2004;103(6):1229-34. PMID: 15172857.
20. American Diabetes Association. Position Statement: Diabetes mellitus. Diabetes Care. 2004;27(Suppl 1):S11-S14. PMID: 14693922.

21. Moses RG, Cheung NW. Point: Universal screening for gestational diabetes mellitus. Diabetes Care. 2009;32(7):1349-51. PMID: 19564479.

22. Danilenko-Dixon DR, Van Winter JT, Nelson RL, et al. Universal versus selective gestational diabetes screening: application of 1997 American Diabetes Association recommendations. Am J Obstet Gynecol. 1999;181(4):798-802. PMID: 10521732.

23. Berger H, Sermer M. Counterpoint: Selective screening for gestational diabetes mellitus. Diabetes Care. 2009;32(7):1352-4. PMID: 19564480.

24. Metzger B, Gabbe S, Persson B, et al. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care. 2010;33(3):676-82. PMID: 20190296.

25. Lipscombe LL, Hux JE. Trends in diabetes prevalence, incidence, and mortality in Ontario, Canada 1995-2005: a population-based study. Lancet. 2007;369(9563):750-6. PMID: 17336651.

26. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 1999;22(Suppl 1):S5-S19.

27. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2000;23(Suppl 1):S4-S19. PMID: 12017675.

28. Report of the expert committee on the diagnosis and classification of diabetes mellitus. Diabetes Care. 2001;24(Suppl.1):S5-S20.

29. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2002;25:S5.

30. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2003;26:S5-20. PMID: 12502614.

31. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2004;27(Suppl 1):S5-S10. PMID: 14693921.

32. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2005;28(Suppl 1):S37-S42. PMID: 15618111.

33. American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2007;30(Suppl 1):S42-S7. PMID: 17192378.

34. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2008;31:S55-S60. PMID: 18165338.

35. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2009;32:S62-S67. PMID: 19118289.

36. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2010;33:S62-S69. PMID: 20042775.

37. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care. 2010;33(3):676-82. PMID: 20190296.

38. Jovanovic L. American Diabetes Association's Fourth International Workshop-Conference on Gestational Diabetes Mellitus: summary and discussion. Therapeutic interventions. Diabetes Care. 1998;21 Suppl 2:B131-B137. PMID: 9704240.

39. Metzger BE, Oats JN, Kjos SL, et al. Summary and Recommendations of the Fifth International Workshop-Conference on Gestational Diabetes Mellitus. Diabetes Care. 2007;30(Suppl 2):s251-s260. PMID: 17596481.

40. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. Diabetes. 1979;28(12):1039-57. PMID: 510803.

41. World Health Organization. Definition, diagnosis and classification of diabetes mellitus and its complications. Report of a WHO Consultation. Part 1: Diagnosis and classification of diabetes mellitus. 1999.

42. World Health Organization. Report of a WHO study Group (Technical Report Series No.727). Report of a WHO study Group (Technical Report Series No.727). 1985.

43. Canadian Diabetes Association Clinical Practice Guidelines Expert Committee. Canadian Diabetes Association 2003 Clinical Practice Guidelines for the Prevention and Management of Diabetes in Canada. Can J Diabetes. 2003;27(suppl 2), S1-S152.

44. Canadian Diabetes Association 2008 Clinical Practice Guidelines for the Prevention and Management of Diabetes in Canada [corrected] [published erratum appears in CAN J DIABETES 2009 Mar;33(1):46]. Canadian Journal of Diabetes. 2008;32:iv.

45. Sempowski IP, Houlden RL. Managing diabetes during pregnancy. Guide for family physicians. Canadian Family Physician Médecin De Famille Canadien. 2003;49:761-7. PMID: 12836864.

46. Metzger BE. Summary and recommendations of the Third International Workshop-Conference on Gestational Diabetes Mellitus. Diabetes. 1991;40 Suppl 2:197-201. PMID: 1748259.

47. Hoffman L, Nolan C, Wilson JD, et al. Gestational diabetes mellitus--management guidelines. The Australasian Diabetes in Pregnancy Society. The Medical Journal Of Australia. 1998;169(2):93-7. PMID: 9700346.

48. Brown CJ, Dawson A, Dodds R, et al. Report of the Pregnancy and Neonatal Care Group. Diabetic Medicine: A Journal Of The British Diabetic Association. 1996;13(9 Suppl 4):S43-S53. PMID: 8894455.

49. American Diabetes Association. Position statement: Gestational diabetes mellitus. Diabetes Care. 2003;26(Suppl 1):S103-S105. PMID: 12502631.

50. Crowther CA, Hiller JE, Moss JR, et al. Effect of treatment of gestational diabetes mellitus on pregnancy outcomes. N Engl J Med. 2005;352(24):2477-86. PMID: 15951574.

51. Langer O, Levy J, Brustman L, et al. Glycemic control in gestational diabetes mellitus--how tight is tight enough: small for gestational age versus large for gestational age? Am J Obstet Gynecol. 1989;161(3):646-53. PMID: 2782347.

52. Cheung NW, Oats JJ, McIntyre HD. Australian carbohydrate intolerance study in pregnant women: implications for the management of gestational diabetes. Aust N Z J Obstet Gynaecol. 2005;45(6):484-5. PMID: 16401212.

53. U.S.Preventive Services Task Force. Screening for gestational diabetes mellitus: recommendations and rationale. Obstet Gynecol. 2003;101(2):393-395. PMID: 12576265.

54. Landon MB, Spong CY, Thom E, et al. A multicenter, randomized trial of treatment for mild gestational diabetes. N Engl J Med. 2009;361(14):1339-48. PMID: 19797280.

55. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36. PMID: 22007046.

56. Rutter C, Gastonis C. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med. 2001;20:2865-84. PMID: 11568945.

57. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58(10):982-90. PMID: 16168343.

58. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-88. PMID: 3802833.

59. Borenstein M, Hedges L, Higgins J, et al. Introduction to meta-analysis. West Sussex, UK: John Wiley & Sons; 2009.

60. Begg C, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics. 1994;50:1088-101. PMID: 7786990.

61. Egger M, Smith G, Schneider M, et al. Bias in meta-analysis detected by a single graphical test. Br Med J. 1997;315(629):634. PMID: 9310563.

62. van LM, Zweers EJ, Opmeer BC, et al. Comparison of accuracy measures of two screening tests for gestational diabetes mellitus. Diabetes Care. 2007;30(11):2779-84. PMID: 17698616.

63. Ayach W, Costa RA, Calderon IM, et al. Comparison between 100-g glucose tolerance test and two other screening tests for gestational diabetes: combined fasting glucose with risk factors and 50-g glucose tolerance test. Rev Paul Med. 2006;124(1):4-9. PMID: 16612455.

64. Yogev Y, Langer O, Xenakis EM, et al. Glucose screening in Mexican-American women. Obstet Gynecol. 2004;103(6):1241-5. PMID: 15172859.

65. Chastang N, Hartemann-Heurtier A, Sachon C, et al. Comparison of two diagnostic tests for gestational diabetes in predicting macrosomia. Diabetes Metab. 2003;29(2 Pt 1):139-44. PMID: 12746634.

66. Perea-Carrasco R, Perez-Coronel R, busac-Aguilar R, et al. A simple index for detection of gestational diabetes mellitus. J R Soc Med. 2002;95(9):435-9. PMID: 12205206.

67. Ardawi MS, Nasrat HA, Jamal HS, et al. Screening for gestational diabetes mellitus in pregnant females. Saudi Med J. 2000;21(2):155-60. PMID: 11533772.

68. Perucchini D, Fischer U, Spinas GA, et al. Using fasting plasma glucose concentrations to screen for gestational diabetes mellitus: prospective population based study. BMJ: British Medical Journal (International Edition). 1999;319(7213):812-5. PMID: 10496823.

69. Lamar ME, Kuehl TJ, Cooney AT, et al. Jelly beans as an alternative to a fifty-gram glucose beverage for gestational diabetes screening. Am J Obstet Gynecol 1999;181(5 Pt 1):1154-7. PMID: 10561636.

70. Siribaddana SH, Deshabandu R, Rajapakse D, et al. The prevalence of gestational diabetes in a Sri Lankan antenatal clinic. Ceylon Med J. 1998;43(2):88-91. PMID: 9704548.

71. Eslamian L, Ramezani Z. Evaluation of a breakfast as screening test for the detection of gestational diabetes. Acta Medica Iranica. 2008;46(1):43-6.

72. Espinosa De Los MA, Parra A, Hidalgo R, et al. The after breakfast 50-g, 1-hour glucose challenge test in urban Mexican pregnant women: Its sensitivity and specificity evaluated by three diagnostic criteria for gestational diabetes mellitus. Acta Obstet Gynecol Scand. 1999;78(4):294-8. PMID: 10203295.

73. Deerochanawong C, Putiyanun C, Wongsuryrat M, et al. Comparison of National Diabetes Data Group and World Health Organization criteria for detecting gestational diabetes mellitus. Diabetologia. 1996;39(9):1070-3. PMID: 8877291.

74. Uncu G, Ozan H, Cengiz C. The comparison of 50 grams glucose challenge test, HbA(1c) and fructosamine levels in diagnosis of gestational diabetes mellitus. Clin Exp Obstet Gynecol. 1995;22(3):230-4. PMID: 7554262.

75. Kashi Z, Borzouei SH, Akha O, et al. Diagnostic value of fasting plasma glucose in screening of gestational diabetes mellitus. Int J Endocrinol Metab. 2007;5(1):1-4.

76. Gandevani SB, Garshasbi A, Dibaj S. Cut-off value of 1-h, 50-g glucose challenge test for screening of gestational diabetes mellitus in an Iranian population. J Obstet Gynaecol Res. 2011;37(6):534-7. PMID: 21375670.

77. Soheilykhah S, Rashidi M, Mojibian M, et al. An appropriate test for diagnosis of gestational diabetes mellitus. Gynecol Endocrinol. 2011;27(10):785-8. PMID: 21250875.

78. Black MH, Sacks DA, Xiang AH, et al. Clinical outcomes of pregnancies complicated by mild gestational diabetes mellitus differ by combinations of abnormal oral glucose tolerance test values. Diabetes Care. 2010;33(12):2524-30. PMID: 20843973.

79. Morikawa M, Yamada T, Yamada T, et al. Change in the number of patients after the adoption of IADPSG criteria for hyperglycemia during pregnancy in Japanese women. Diabetes Res Clin Pract. 2010;90(3):339-42. PMID: 20870307.

80. Corrado F, Benedetto AD, Cannata ML, et al. A single abnormal value of the glucose tolerance test is related to increased adverse perinatal outcome. J Matern Fetal Neonatal Med. 2009;22(7):597-601. PMID: 19488948.

81. Cheng YW, Block-Kurbisch I, Caughey AB. Carpenter-Coustan criteria compared with the national diabetes data group thresholds for gestational diabetes mellitus. Obstet Gynecol. 2009;114(2:Pt 1):326-32. PMID: 19622994.

82. Biri A, Korucuoglu U, Ozcan P, et al. Effect of different degrees of glucose intolerance on maternal and perinatal outcomes. J Matern Fetal Neonatal Med. 2009;22(6):473-8. PMID: 19479645.

83. Retnakaran R, Qi Y, Sermer M, et al. Glucose intolerance in pregnancy and future risk of pre-diabetes or diabetes. Diabetes Care. 2008;31(10):2026-31. PMID: 18628572.

84. Shirazian N, Mahboubi M, Emdadi R, et al. Comparison of different diagnostic criteria for gestational diabetes mellitus based on the 75-g oral glucose tolerance test: a cohort study. Endocrine Pract. 2008;14(3):312-7. PMID: 18463038.

85. Kwik M, Seeho SK, Smith C, et al. Outcomes of pregnancies affected by impaired glucose tolerance. Diabetes Res Clin Pract. 2007;77(2):263-8. PMID: 17275121.

86. Chico A, Lopez-Rodo V, Rodriguez-Vaca D, et al. Features and outcome of pregnancies complicated by impaired glucose tolerance and gestational diabetes diagnosed using different criteria in a Spanish population. Diabetes Res Clin Pract. 2005;68(2):141-6. PMID: 15860242.

87. Bo S, Menato G, Gallo ML, et al. Mild gestational hyperglycemia, the metabolic syndrome and adverse neonatal outcomes. Acta Obstet Gynecol Scand. 2004;83(4):335-40. PMID: 15005779.

88. Stamilio DM, Olsen T, Ratcliffe S, et al. False-positive 1-hour glucose challenge test and adverse perinatal outcomes. Obstet Gynecol. 2004;103(1):148-56. PMID: 14704259.

89. Pennison EH, Egerman RS. Perinatal outcomes in gestational diabetes: a comparison of criteria for diagnosis. Am J Obstet Gynecol. 2001;184(6):1118-21. PMID: 11349174.

90. Berggren EK, Boggess KA, Stuebe AM, et al. National Diabetes Data Group vs. Carpenter-Coustan criteria to diagnose gestational diabetes. Am J Obstet Gynecol. 2011;205(3):253-e1-7. PMID: 22071053.

91. Sacks DA, Greenspoon JS, bu-Fadil S, et al. Toward universal criteria for gestational diabetes: The 75-gram glucose tolerance test in pregnancy. Am J Obstet Gynecol. 1995;172(2 I):607-14. PMID: 7856693.

92. Chou C, Lin C, Yang C, et al. Pregnancy outcomes of Taiwanese women with gestational diabetes mellitus: a comparison of Carpenter-Coustan and National Diabetes Data Group criteria. J Womens Health. 2010;19(5):935-8. PMID: 20370431.

93. Lapolla A, Dalfra MG, Ragazzi E, et al. New International Association of the Diabetes and Pregnancy Study Groups (IADPSG) recommendations for diagnosing gestational diabetes compared with former criteria: a retrospective study on pregnancy outcome. Diabetic Med. 2011;28(9):1074-7. PMID: 21658125.

94. Cok T, Tarim E, Bagis T. Isolated abnormal value during the 3-hour glucose tolerance test: which value is associated with macrosomia? J Matern Fetal Neonatal Med. 2011;24(8):1039-41. PMID: 21247232.

95. Fassett MJ, Dhillon SH, Williams TR. Effects on perinatal outcome of treating women with 1 elevated glucose tolerance test value. Am J Obstet Gynecol. 2007;196(6):597.e1-4. PMID: 17547912.

96. Malcolm JC, Lawson ML, Gaboury I, et al. Glucose tolerance of offspring of mother with gestational diabetes mellitus in a low-risk population. Diabetic Med. 2006;23(5):565-70. PMID: 16681566.

97. Bonomo M, Corica D, Mion E, et al. Evaluating the therapeutic approach in pregnancies complicated by borderline glucose intolerance: a randomized clinical trial. Diabetic Med. 2005;22(11):1536-41. PMID: 16241919.

98. Bevier WC, 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. Am J Perinatol. 1999;16(6):269-75. PMID: 10586979.

99. Agarwal MM, Hughes PF, Punnose J, et al. Fasting plasma glucose as a screening test for gestational diabetes in a multi-ethnic, high-risk population. Diabet Med. 2000;17(10):720-6. PMID: 11110505.

100. Agarwal MM, Hughes PF, Punnose J, et al. Gestational diabetes screening of a multiethnic, high-risk population using glycated proteins. Diabetes Research and Clinical Practice. 2001;51(1):67-73. PMID: 11137184.

101. Maegawa Y, Sugiyama T, Kusaka H, et al. Screening tests for gestational diabetes in Japan in the 1st and 2nd trimester of pregnancy. Diabetes Research and Clinical Practice. 2003;62(1):47-53. PMID: 14581157.

102. Ricart W, Lopez J, Mozas J, et al. Potential impact of American Diabetes Association (2000) criteria for diagnosis of gestational diabetes mellitus in Spain. Diabetologia. 2005;48(6):1135-41. PMID: 15889233.

103. Kim HS, Chang KH, Yang JI, et al. Clinical outcomes of pregnancy with one elevated glucose tolerance test value. Int J Gynaecol Obstet. 2002;78(2):131-8. PMID: 12175714.

104. Bobrowski RA, Bottoms SF, Micallef JA, et al. Is the 50-gram glucose screening test ever diagnostic? J Matern Fetal Med. 1996;5(6):317-20. PMID: 8972407.

105. Rey E, Hudon L, Michon N, et al. Fasting plasma glucose versus glucose challenge test: screening for gestational diabetes and cost effectiveness. Clin Biochem. 2004;37(9):780-4. PMID: 15329316.

106. Vambergue A, Nuttens MC, Verier-Mine O, et al. Is mild gestational hyperglycaemia associated with maternal and neonatal complications? The Diagest Study. Diabet Med. 2000;17(3):203-8. PMID: 10784224.

107. Rajput R, Yogeshyadav, Rajput M, et al. Utility of HbA(1c) for diagnosis of gestational diabetes mellitus. Diabetes Res Clin Pract. 2012. PMID: 22456454.

108. Chevalier N, Fenichel P, Giaume V, et al. Universal two-step screening strategy for gestational diabetes has weak relevance in French Mediterranean women: should we simplify the screening strategy for gestational diabetes in France? Diabetes Metab. 2011;37(5):419-25. PMID: 21489844.

109. Agarwal MM, Dhatt GS, Othman Y, et al. Gestational diabetes: an evaluation of serum fructosamine as a screening test in a high-risk population. Gynecol Obstet Invest. 2011;71(3):207-12. PMID: 21160150.

110. Agarwal MM, Dhatt GS, Safraou MF. Gestational diabetes: using a portable glucometer to simplify the approach to screening. Gynecol Obstet Invest. 2008;66(3):178-83. PMID: 18562798.

111. Wijeyaratne CN, Ginige S, Arasalingam A, et al. Screening for gestational diabetes mellitus: the Sri Lankan experience. Ceylon Med J. 2006;51(2):53-8. PMID: 17180809.

112. Agarwal MM, Dhatt GS, Punnose J. Gestational diabetes: utility of fasting plasma glucose as a screening test depends on the diagnostic criteria. Diabetic Med. 2006;23(12):1319-26. PMID: 17116182.

113. Agarwal MM, Dhatt GS, Punnose J, et al. Gestational diabetes: a reappraisal of HBA1c as a screening test. Acta Obstet Gynecol Scand. 2005;84(12):1159-63. PMID: 16305701.

114. Hill JC, Krishnaveni GV, Annamma I, et al. Glucose tolerance in pregnancy in South India: relationships to neonatal anthropometry. Acta Obstet Gynecol Scand. 2005;84(2):159-65. PMID: 15683377.

115. Jensen DM, Molsted-Pedersen L, Beck-Nielsen H, et al. Screening for gestational diabetes mellitus by a model based on risk indicators: a prospective study. Am J Obstet Gynecol. 2003;189(5):1383-8. PMID: 14634573.

116. Buhling KJ, Henrich W, Kjos SL, et al. Comparison of point-of-care-testing glucose meters with standard laboratory measurement of the 50g-glucose-challenge test (GCT) during pregnancy. Clin Biochem. 2003;36(5):333-7. PMID: 12849863.

117. Jakobi P, Weissman A, Egozi J, et al. Perinatal significance of diagnosing glucose intolerance during pregnancy with portable glucose meter. J Perinat Med. 2003;31(2):140-5. PMID: 12747230.

118. Soonthornpun S, Soonthornpun K, Aksonteing J, et al. A comparison between a 75-g and 100-g oral glucose tolerance test in pregnant women. Int J Gynaecol Obstet. 2003;81(2):169-73. PMID: 12706274.

119. Ostlund I, Hanson U. Occurrence of gestational diabetes mellitus and the value of different screening indicators for the oral glucose tolerance test. Acta Obstet Gynecol Scand. 2003;82(2):103-8. PMID: 12648169.

120. Reichelt AJ, Spichler ER, Branchtein L, et al. Fasting plasma glucose is a useful test for the detection of gestational diabetes. Brazilian Study of Gestational Diabetes (EBDG) Working Group. Diabetes Care. 1998;21(8):1246-9. PMID: 9702428.

121. Rust O, Bofill JA, Carroll SC, et al. Two-hour postprandial test versus one-hour, fifty-gram glucola test as screening tools for gestational diabetes: a critical analysis. J Perinatol. 1998;18(1):49-54. PMID: 9527945.

122. Fung HY, Wong SP, Rogers M. The influence of glucose tolerance tests on subsequent carbohydrate metabolism in pregnancy. Acta Obstet Gynecol Scand. 1996;75(4):347-51. PMID: 8638454.

123. Berkus MD, Langer O. Glucose tolerance test periodicity: the effect of glucose loading. Obstet Gynecol. 1995;85(3):423-7. PMID: 7862384.

124. Moses RG, Morris GJ, Petocz P, et al. The impact of potential new diagnostic criteria on the prevalence of gestational diabetes mellitus in Australia. The Medical Journal Of Australia. 2011;194(7):338-40. PMID: 21470082.

125. Buhling KJ, Elze L, Henrich W, et al. The usefulness of glycosuria and the influence of maternal blood pressure in screening for gestational diabetes. Eur J Obstet Gynecol Reprod Biol. 2004;113(2):145-8. PMID: 15063950.

126. Sacks DA, Chen W, Wolde-Tsadik G, et al. Fasting plasma glucose test at the first prenatal visit as a screen for gestational diabetes. Obstet Gynecol. 2003;101(6):1197-203. PMID: 12798525.

127. Kauffman RP, Castracane VD, Peghee D, et al. Detection of gestational diabetes mellitus by homeostatic indices of insulin sensitivity: A preliminary study. Am J Obstet Gynecol. 2006;194(6):1576-82. PMID: 16638603.

128. Balaji V, Madhuri BS, Paneerselvam A, et al. Comparison of Venous Plasma Glucose and Capillary Whole Blood Glucose in the Diagnosis of Gestational Diabetes Mellitus: A Community-Based Study. Diabetes Technol Ther. 2011;14(2):131-4. PMID: 21992269.

129. Balaji V, Balaji M, Anjalakshi C, et al. Inadequacy of fasting plasma glucose to diagnose gestational diabetes mellitus in Asian Indian women. Diabetes Res Clin Pract. 2011;94(1):e21-e23. PMID: 21831468.

130. Chanprapaph P, Sutjarit C. Prevalence of gestational diabetes mellitus (GDM) in women screened by glucose challenge test (GCT) at Maharaj Nakorn Chiang Mai Hospital. J Med Assoc Thai. 2004;87(10):1141-6. PMID: 15560687.

131. Solomon CG, Willett WC, Rich-Edwards J, et al. Variability in diagnostic evaluation and criteria for gestational diabetes. Diabetes Care. 1996;19(1):12-6. PMID: 8720526.

132. Hillier TA, Pedula KL, Schmidt MM, et al. Childhood obesity and metabolic imprinting: the ongoing effects of maternal hyperglycemia. Diabetes Care. 2007;30(9):2287-92. PMID: 17519427.

133. Jensen DM, Damm P, Sorensen B, et al. Proposed diagnostic thresholds for gestational diabetes mellitus according to a 75-g oral glucose tolerance test. Maternal and perinatal outcomes in 3260 Danish women. Diabetic Med. 2003;20(1):51-7. PMID: 12519320.

134. Lao TT, Tam KF. Gestational diabetes diagnosed in third trimester pregnancy and pregnancy outcome. Acta Obstet Gynecol Scand. 2001;80(11):1003-8. PMID: 11703196.

135. Rust OA, Bofill JA, Andrew ME, et al. Lowering the threshold for the diagnosis of gestational diabetes. Am J Obstet Gynecol. 1996;175(4 Pt 1):961-5. PMID: 8885755.

136. Berkus MD, Langer O, Piper JM, et al. Efficiency of lower threshold criteria for the diagnosis of gestational diabetes. Obstet Gynecol. 1995;86(6):892-6. PMID: 7501334.

137. Tan PC, Ling LP, Omar SZ. The 50-g glucose challenge test and pregnancy outcome in a multiethnic Asian population at high risk for gestational diabetes. Int J Gynaecol Obstet. 2009;105(1):50-5. PMID: 19154997.

138. Agarwal MM, Dhatt GS, Punnose J, et al. Gestational diabetes in a high-risk population: using the fasting plasma glucose to simplify the diagnostic algorithm. Eur J Obstet Gynecol Reprod Biol. 2005;120(1):39-44. PMID: 15866084.

139. Yachi Y, Tanaka Y, Anasako Y, et al. Contribution of first trimester fasting plasma insulin levels to the incidence of glucose intolerance in later pregnancy: Tanaka women's clinic study. Diabetes Res Clin Pract. 2011;92(2):293-8. PMID: 21396732.

140. Weerakiet S, Lertnarkorn K, Panburana P, et al. Can adiponectin predict gestational diabetes? Gynecol Endocrinol. 2006;22(7):362-8. PMID: 16864145.

141. Bito T, Nyari T, Kovacs L, et al. 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. Eur J Obstet Gynecol Reprod Biol. 2005;121(1):51-5. PMID: 15989984.

142. Sermer M, Naylor CD, Gare DJ, et al. Impact of increasing carbohydrate intolerance on maternal-fetal outcomes in 3637 women without gestational diabetes. The Toronto Tri-Hospital Gestational Diabetes Project. Am J Obstet Gynecol. 1995;173(1):146-56. PMID: 7631672.

143. Brustman LE, Gela BD, Moore M, et al. Variations in oral glucose tolerance tests: the 100- versus 75-g controversy. J Assoc Acad Minor Phys. 1995;6(2):70-2. PMID: 7772935.

144. Cetin M, Cetin A. Time-dependent gestational diabetes screening values. Int J Gynecol Obstet. 1997;56(3):257-61. PMID: 9127158.

145. Lapolla A, Dalfra MG, Bonomo M, et al. Can plasma glucose and HbA1c predict fetal growth in mothers with different glucose tolerance levels? Diabetes Res Clin Pract. 2007;77(3):465-70. PMID: 17350135.

146. Langer O, Yogev Y, Most O, et al. Gestational diabetes: the consequences of not treating. Am J Obstet Gynecol. 2005;192(4):989-97. PMID: 15846171.

147. Lao TT, Ho LF. Does maternal glucose intolerance affect the length of gestation in singleton pregnancies? J Soc Gynecol Investig. 2003;10(6):366-71. PMID: 12969780.

148. Bonomo M, Gandini ML, Farina A, et al. Should we treat minor degrees of glucose intolerance in pregnancy? Ann Ist Super Sanita. 1997;33(3):393-7. PMID: 9542269.

149. Nord E, Hanson U, Persson B. Blood glucose limits in the diagnosis of impaired glucose tolerance during pregnancy. Relation to morbidity. Acta Obstet Gynecol Scand. 1995;74(8):589-93. PMID: 7660761.

150. Schwartz ML, Ray WN, Lubarsky SL, et al. The diagnosis and classification of gestational diabetes mellitus: Is it time to change our tune? Am J Obstet Gynecol. 1999;180(6 I):1560-71. PMID: 10368504.

151. Poyhonen-Alho MK, Teramo KA, Kaaja RJ, et al. 50gram oral glucose challenge test combined with risk factor-based screening for gestational diabetes. Eur J Obstet Gynecol Reprod Biol. 2005;121(1):34-7. PMID: 15989983.

152. Adams KM, Li H, Nelson RL, et al. Sequelae of unrecognized gestational diabetes. Am J Obstet Gynecol. 1998;178(6):1321-32. PMID: 9662318.

153. Mello G, Elena P, Ognibene A, et al. Lack of concordance between the 75-g and 100-g glucose load tests for the diagnosis of gestational diabetes mellitus. Clin Chem. 2006;52(9):1679-84. PMID: 16873295.

154. Aberg A, Rydhstroem H, Frid A. Impaired glucose tolerance associated with adverse pregnancy outcome: a population-based study in southern Sweden. Am J Obstet Gynecol. 2001;184(2):77-83. PMID: 11174484.

155. Yang X, Hsu-Hage B, Zhang H, et al. Women with impaired glucose tolerance during pregnancy have significantly poor pregnancy outcomes. Diabetes Care. 2002;25(9):1619-24. PMID: 12196437.

156. Sun B, Wang X, Song Q, et al. Prospective studies on the relationship between the 50 g glucose challenge test and pregnant outcome. Chin Med J. 1995;108(12):910-3. PMID: 8728943.

157. Tan PC, Ling LP, Omar SZ. Screening for gestational diabetes at antenatal booking in a Malaysian university hospital: the role of risk factors and threshold value for the 50-g glucose challenge test. Aust N Z J Obstet Gynaecol. 2007;47(3):191-7. PMID: 17550485.

158. O'Sullivan JB, Mahan CM. Criteria for the oral glucose tolerance test in pregnancy. Diabetes. 1964;13:278-85. PMID: 14166677.

159. Sermer M, Naylor CD, Farine D, et al. The Toronto Tri-Hospital Gestational Diabetes Project. A preliminary review. Diabetes Care. 1998;21(Suppl 2):B33-B42. PMID: 9704225.

160. Naylor CD, Sermer M, Chen E, et al. Cesarean delivery in relation to birth weight and gestational glucose tolerance: pathophysiology or practice style? Toronto Trihospital Gestational Diabetes Investigators. JAMA. 1996;275(15):1165-70. PMID: 8609683.

161. Eslamian L, Ramezani Z. Breakfast as a screening test for gestational diabetes. Int J Gynaecol Obstet. 2007;96(1):34-5. PMID: 17188692.

162. van LM, Opmeer BC, Zweers EJ, et al. Estimating the risk of gestational diabetes mellitus: a clinical prediction model based on patient characteristics and medical history. BJOG. 2010;117(1):69-75. PMID: 20002371.

163. Landon MB, Mele L, Spong CY, et al. The relationship between maternal glycemia and perinatal outcome. Obstet Gynecol. 2011;117(2 Pt 1):218-24. PMID: 21309194.

164. Pirc LK, Owens JA, Crowther CA, et al. Mild gestational diabetes in pregnancy and the adipoinsular axis in babies born to mothers in the ACHOIS randomised controlled trial. BMC Pediatr. 2007;7:18. PMID: 17430602.

165. Rasmussen SS, Glumer C, Sandbaek A, et al. Short-term reproducibility of impaired fasting glycaemia, impaired glucose tolerance and diabetes The ADDITION study, DK. Diabetes Res Clin Pract. 2008;80(1):146-52. PMID: 18082284.

166. Hyperglycemia and adverse pregnency outcome (HAPO) study Cooperative Research Group. Hyperglycemia and adverse pregnancy outcomes. The New England Journal of Medicine. 2008;358(19):1991-2002. PMID: 18463375.

167. Naylor CD, Sermer M, Chen EL, et al. Selective screening for gestational diabetes mellitus. N Engl J Med. 1997;337(22):1591-6. PMID: 9371855.

168. Phillips DI. Programming of the stress response: a fundamental mechanism underlying the long-term effects of the fetal environment? J Intern Med. 2007;261(5):453-60. PMID: 17444884.

169. Daniells S, Grenyer BFS, Davis WS, et al. Gestational diabetes mellitus: Is a diagnosis associated with an increase in maternal anxiety and stress in the short and intermediate term? Diabetes Care. 2003;26(2):385-9. PMID: 12547867.

170. Horvath K, Koch K, Jeitler K, et al. Effects of treatment in women with gestational diabetes mellitus: systematic review and meta-analysis. BMJ: British Medical Journal (International Edition). 2010;340:c1395. PMID: 20360215.

171. Pettitt DJ, McKenna S, McLaughlin C, et al. Maternal glucose at 28 weeks of gestation is not associated with obesity in 2-year-old offspring: The Belfast Hyperglycemia and Adverse Pregnancy Outcome (HAPO) family study. Diabetes Care. 2010;33(6):1219-23. PMID: 20215449.

172. Ryan EA. Diagnosing gestational diabetes. Diabetologia 2011;54(3):480-6. PMID: 21203743.

173. HAPO Study Cooperative Research Group. Hyperglycaemia and Adverse Pregnancy Outcome (HAPO) Study: associations with maternal body mass index. BJOG. 2010;117(5):575-84. PMID: 20089115.

174. Ricart W, Lopez J, Mozas J, et al. Body mass index has a greater impact on pregnancy outcomes than gestational hyperglycaemia. Diabetologia. 2005;48(9):1736-42. PMID: 16052327.

175. Cundy T, Gamble G, Townend K, et al. Perinatal mortality in Type 2 diabetes mellitus. Diabet Med. 2000;17(1):33-9. PMID: 10691157.

176. Schaefer-Graf UM, Hartmann R, Pawliczak J, et al. Association of breast-feeding and early childhood overweight in children from mothers with gestational diabetes mellitus. Diabetes Care. 2006;29(5):1105-7. PMID: 16644645.

177. Buchanan TA, Kjos SL, Montoro MN, et al. Use of fetal ultrasound to select metabolic therapy for pregnancies complicated by mild gestational diabetes. Diabetes Care. 1994;17(4):275-83. PMID: 8026282.

Acronyms and Abbreviations

ACHOIS	Australian Carbohydrate Intolerance in Pregnant Women Study
ACOG	American Congress of Obstetricians and Gynecologists
ADA	American Diabetes Association
ADIPS	Australasian Diabetes in Pregnancy Society
BMI	Body-mass index
CC	Carpenter and Coustan
CI	Confidence interval
D	Day(s)
dL	Deciliter
DM	Diabetes mellitus
Dx	Diagnosis/diagnostic
EASD	European Association for the Study of Diabetes
FPG	Fasting plasma glucose
GCT/OGCT	Glucose tolerance test and oral glucose tolerance test are synonymous
GDM	Gestational diabetes mellitus
g(s)	Gram(s)
h(s)	Hour(s)
HSROC	Hierarchical summary receiver operator characteristic
HAPO	Hyperglycemia and Adverse Pregnancy Outcomes Study
HbA1c	Glycated Hemoglobin, Hemoglobin A1c
IFG	Impaired fasting glucose
IGT	Impaired glucose tolerance
IGT-2	Double impaired glucose tolerance
IADPSG	International Association of the Diabetes in Pregnancy Study Groups
IQR	Inter-quartile range
IWC	International Workshop Conference
JSOG	Japan Society of Obstetrics and Gynecology
kg	kilogram
LGA	Large for gestational age
L	Liter
m	Meter
MD	Mean difference
µmol	Micromole
mg	Milligrams
mmol	Millimole
mo(s)	Month(s)
NDDG	National Diabetes Data Group
NICU	Neonatal Intensive Care Unit
NOS	Newcastle-Ottawa Quality Assessment Scale
NR	Not reported
N or n	Number

NPV	Negative predictive value
OGCT	Oral glucose challenge text
OGTT	Oral glucose tolerance test
PCS	Prospective cohort study
PPV	Positive predictive value
QUADAS	Quality assessment of diagnostic accuracy studies
RCS	Retrospective cohort study
RCT(s)	Randomized controlled trial(s)
RDS	Respiratory distress syndrome
RR	Risk ratio (or relative risk)
Sn	Sensitivity
Sp	Specificity
SD	Standard deviation
SGA	Small for gestational age
WHO	World Health Organization
wk(s)	Week(s)
yr(s)	Year(s)

Appendix A. Literature Search Strings

Table A1. Medline

Database: Medline via Ovid <1948 to September Week 4 2011>
Search Date: 9 October 2011
Results: 8,234

1. Diabetes, Gestational/
2. Fetal Macrosomia/
3. Pregnancy Complications/
4. GDM.tw.
5. (gestation$ adj2 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
6. (pregnan$ adj3 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
7. (maternal adj2 (diabet$ or DM or glyc?emia or hyperglyc?emia)).tw.
8. (hyperglyc?emia adj2 pregnan$).tw.
9. macrosomia.tw.
10. or/1-9
11. mass screening/
12. prenatal diagnosis/
13. screen$.tw.
14. ((prenatal or early) adj2 diagnosis).tw.
15. Glucose Tolerance Test/
16. Glucose Intolerance/
17. Blood Glucose/
18. Risk Factors/
19. (glucose adj (tolerance or intolerance or challenge)).tw.
20. OGTT.tw.
21. GCT.tw.
22. (fasting adj2 glucose).tw.
23. or/11-22
24. "Sensitivity and Specificity"/
25. "Predictive Value of Tests"/
26. ROC Curve/
27. specific$.tw.
28. sensitiv$.tw.
29. predictive value.tw.
30. accurac$.tw.
31. diagnostic errors/
32. diagnostic error?.tw.
33. false negative reactions/
34. false positive reactions/
35. (false adj (negative or positive)).tw.
36. "reproducibility of results"/
37. reference values/
38. reference standards/
39. or/24-38
40. and/10,23,39
41. intervention?.mp.
42. (treating or treatment? or therapy or therapies).mp.
43. manage$.mp.
44. monitor$.mp.
45. exp sulfonylurea compounds/
46. Gliclazide/
47. Glyburide/
48. Tolbutamide/
49. sulfonylurea?.tw.
50. gliclazid$.tw.
51. glimepirid$.tw.
52. glipizid$.tw.
53. glyburid$.tw.
54. tolbutamid$.tw.
55. (antidiabet$ or anti-diabet$).tw.
56. insulin?.mp.
57. glibenclamid$.mp.
58. acarbos$.mp.

59. exp Diet Therapy/
60. (diet adj2 (therap$ or restrict$ or advice)).tw.
61. medical nutrition$ therapy.tw.
62. MNT.tw.
63. exp Life Style/
64. (lifestyle$ or life-style$).mp.
65. Blood Glucose Self-Monitoring/
66. (blood glucose adj (self monitor$ or self-monitor$)).tw.
67. ((self monitor$ or self-monitor$) adj blood glucose).tw.
68. SMBG.tw.
69. Counseling/
70. counsel$.tw.
71. Labor, Induced/
72. (induc$ adj2 labo?r).tw.
73. exp Cesarean Section/
74. c?esarean.tw.
75. exp Pregnancy Outcome/
76. pregnanc$ outcome?.tw.
77. or/41-76
78. and/10,77
79. or/40,78
80. clinical trial.pt.
81. randomized controlled trial.pt.
82. randomi?ed.ti,ab.
83. placebo.ti,ab.
84. dt.fs.
85. randomly.ti,ab.
86. trial.ti,ab.
87. groups.ti,ab.
88. or/80-87
89. animals/
90. humans/
91. 89 not (89 and 90)
92. 88 not 91
93. cohort studies/
94. follow-up studies/
95. longitudinal studies/
96. prospective studies/
97. retrospective studies/
98. ((cohort? or follow-up or followup or longitud$ or prospectiv$ or retrospectiv$) adj (study or studies or trial?)).tw.
99. or/93-98
100. 99 not 91
101. exp Guideline/
102. Health Planning Guidelines/
103. (clinical adj2 guideline?).tw.
104. CPG?.tw.
105. ((practice or consensus or position) adj2 (guideline? or recommendation? or statement?)).tw.
106. standard?.tw.
107. protocol?.tw.
108. or/101-107
109. meta analysis.mp,pt.
110. review.pt.
111. search:.tw.
112. or/109-111 [Reviews balanced - HIRU]
113. and/79,92 [Clinical trials & RCTs]
114. and/79,100 [Observational studies]
115. and/79,108 [Guidelines]
116. and/79,112 [SRs MAs]
117. or/113-116
118. limit 117 to (english language and yr="2000 -Current")
119. limit 117 to (english language and yr="2000 -2005")
120. limit 117 to (english language and yr="2006 -Current")

121. remove duplicates from 119
122. remove duplicates from 120
123. or/121-122
124. 113 or 114 or 115
125. 113 or 114 or 115
126. limit 125 to (english language and yr="2000 -Current")
127. limit 125 to (english language and yr="2000 -2005")
128. remove duplicates from 127
129. limit 125 to (english language and yr="2006 -Current")
130. remove duplicates from 129
131. 128 or 130
132. 113 or 114
133. limit 132 to (english language and yr="2000 -Current")
134. limit 132 to (english language and yr="2000 -2005")
135. remove duplicates from 134
136. limit 132 to (english language and yr="2006 -Current")
137. remove duplicates from 136
138. 135 or 137

Table A2. Embase

Database: Embase via Ovid <1996 to 2011 Week 40>
Search Date: 10 October 2011
Results: 5,188

1. pregnancy diabetes mellitus/
2. maternal diabetes mellitus/
3. pregnancy complication/
4. macrosomia/
5. GDM.tw.
6. (gestation$ adj2 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
7. (pregnan$ adj3 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
8. (maternal adj2 (diabet$ or DM or glyc?emia or hyperglyc?emia)).mp.
9. (hyperglyc?emia adj2 pregnan$).tw.
10. macrosomia.tw.
11. or/1-10
12. prenatal screening/
13. early diagnosis/
14. screen$.tw.
15. ((prenatal or early) adj2 diagnosis).tw.
16. exp glucose tolerance test/
17. glucose intolerance/
18. glucose blood level/
19. risk factor/
20. (glucose adj (tolerance or intolerance or challenge)).tw.
21. OGTT.tw.
22. GCT.tw.
23. (fasting adj2 glucose).tw.
24. or/12-23
25. "sensitivity and specificity"/
26. predictive value/
27. receiver operating characteristic/
28. specific$.tw.
29. sensitiv$.tw.
30. predictive value.tw.
31. accurac$.tw.
32. diagnostic error/
33. diagnostic accuracy/
34. diagnostic error?.tw.
35. false negative result/
36. false positive result/
37. (false adj (negative or positive)).tw.
38. reproducibility/

39. reference value/
40. standard/
41. or/25-40
42. and/11,24,41
43. intervention?.mp.
44. (treating or treatment? or therapy or therapies).mp.
45. manage$.mp.
46. monitor$.mp.
47. sulfonylurea derivative/
48. gliclazide/
49. glibenclamide/
50. glimepiride/
51. glipizide/
52. tolbutamide/
53. sulfonylurea?.tw.
54. gliclazid$.tw.
55. glimepirid$.tw.
56. glipizid$.tw.
57. glyburid$.tw.
58. tolbutamid$.tw.
59. (antidiabet$ or anti-diabet$).tw.
60. insulin?.mp.
61. glibenclamid$.mp.
62. acarbos$.mp.
63. exp diet therapy/
64. (diet adj2 (therap$ or restrict$ or advice)).tw.
65. medical nutrition$ therapy.tw.
66. MNT.tw.
67. exp lifestyle/
68. (lifestyle$ or life-style$).mp.
69. blood glucose monitoring/
70. (blood glucose adj (self monitor$ or self-monitor$)).tw.
71. ((self monitior$ or self-monitor$) adj blood glucose).tw.
72. SMBG.tw.
73. counseling/
74. nutritional counseling/
75. counsel$.tw.
76. labor induction/
77. (induc$ adj2 labo?r).tw.
78. cesarean section/
79. c?esarean.tw.
80. pregnancy outcome/
81. pregnanc$ outcome?.tw.
82. or/43-81
83. and/11,82
84. or/42,83
85. clinical trial/
86. randomized controlled trial/
87. randomization/
88. single blind procedure/
89. double blind procedure/
90. crossover procedure/
91. placebo/
92. randomi?ed controlled trial?.tw.
93. RCT.tw.
94. random allocation.tw.
95. randomly allocated.tw.
96. allocated randomly.tw.
97. (allocated adj2 random).tw.
98. single blind$.tw.
99. double blind$.tw.
100. ((treble or triple) adj blind$).tw.

101. placebo$.tw.
102. prospective study/
103. or/85-102
104. case study/
105. case report.tw.
106. abstract report/ or letter/
107. or/104-106
108. 103 not 107 [SIGN Embase RCT filter]
109. animal/
110. human/
111. 109 not (109 and 110)
112. 108 not 111
113. cohort analysis/
114. follow up/
115. longitudinal study/
116. prospective study/
117. retrospective study/
118. ((cohort? or follow-up or followup or longitud$ or prospectiv$ or retrospectiv$) adj (study or studies or trial?)).tw.
119. or/113-118
120. 119 not 111
121. exp practice guideline/
122. (clinical adj2 guideline?).tw.
123. CPG?.tw.
124. ((practice or consensus or position) adj2 (guideline? or recommendation? or statement?)).tw.
125. standard?.tw.
126. protocol?.tw.
127. or/121-126 [Guidelines]
128. and/84,112 [RCTs]
129. and/84,120 [Observational studies]
130. and/84,127 [Guidelines]
131. or/128-130
132. limit 131 to (english language and yr="2000 -2005")
133. remove duplicates from 132
134. limit 131 to (english language and yr="2006 -Current")
135. remove duplicates from 134

133 or 135

Table A3. EMB Reviews

Databases:
Cochrane Central Register of Controlled Trials (CCTR) via Ovid <3rd Quarter 2011>
Cochrane Database of Systematic Reviews (CDSR) via Ovid <2005 to September 2011>
Database of Abstracts of Reviews of Effects (DARE) via Ovid <3rd Quarter 2011>
Search Date: 9 October 2011
Results: CCTR: 23; CDSR: 79; DARE: 23

1. GDM.tw.
2. (gestation$ adj2 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
3. (pregnan$ adj3 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
4. (maternal adj2 (diabet$ or DM or glyc?emia or hyperglyc?emia)).tw.
5. (hyperglyc?emia adj2 pregnan$).tw.
6. macrosomia.tw.
7. or/1-6
8. screen$.tw.
9. ((prenatal or early) adj2 diagnosis).tw.
10. blood glucose.tw.
11. risk factor?.tw.
12. (glucose adj (tolerance or intolerance or challenge)).tw.
13. OGTT.tw.
14. GCT.tw.
15. (fasting adj2 glucose).tw.
16. or/8-15
17. specific$.tw.

18. sensitiv$.tw.
19. predictive value.tw.
20. (ROC or "receiver operating characteristic?").tw.
21. accurac$.tw.
22. diagnostic error?.tw.
23. (false adj (negative or positive)).tw.
24. "reproducibility of results".tw.
25. (reference adj2 (standard? or value?)).tw.
26. or/17-25
27. and/7,16,26
28. intervention?.mp.
29. (treating or treatment? or therapy or therapies).mp.
30. manage$.mp.
31. monitor$.mp.
32. sulfonylurea?.tw.
33. gliclazid$.tw.
34. glimepirid$.tw.
35. glipizid$.tw.
36. glyburid$.tw.
37. tolbutamid$.tw.
38. (antidiabet$ or anti-diabet$).tw.
39. insulin?.mp.
40. glibenclamid$.mp.
41. acarbos$.mp.
42. (diet adj2 (therap$ or restrict$ or advice)).tw.
43. medical nutrition$ therapy.tw.
44. MNT.tw.
45. (lifestyle$ or life-style$).mp.
46. (blood glucose adj (self monitor$ or self-monitor$)).tw.
47. ((self monitior$ or self-monitor$) adj blood glucose).tw.
48. SMBG.tw.
49. counsel$.tw.
50. (induc$ adj2 labo?r).tw.
51. c?esarean.tw.
52. pregnanc$ outcome?.tw.
53. or/28-52
54. and/7,53
55. or/27,54
56. clinical trial.pt.
57. randomized controlled trial.pt.
58. randomi?ed.ti,ab.
59. placebo.ti,ab.
60. dt.fs.
61. randomly.ti,ab.
62. trial.ti,ab.
63. groups.ti,ab.
64. or/56-63
65. (animal? not (animal? and human?)).mp.
66. 64 not 65
67. ((cohort? or follow-up or followup or longitud$ or prospectiv$ or retrospectiv$) adj (study or studies or trial?)).tw.
68. 67 not 66
69. (clinical adj2 guideline?).tw.
70. CPG?.tw.
71. ((practice or consensus or position) adj2 (guideline? or recommendation? or statement?)).tw.
72. standard?.tw.
73. protocol?.tw.
74. or/69-73
75. and/55,66 [Clinical trials & RCTs]
76. and/55,68 [Observational studies]
77. and/55,74 [Guidelines]
78. or/75-77
79. limit 78 to (english language and yr="2000-Current")

80. remove duplicates from 79

Table A4. Global Health

Database: Global Health via Ovid <1973 to September 2011>
Search Date: 9 October 2011
Results: 361

1. GDM.tw.
2. (gestation$ adj2 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
3. (pregnan$ adj3 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
4. (maternal adj2 (diabet$ or DM or glyc?emia or hyperglyc?emia)).tw.
5. (hyperglyc?emia adj2 pregnan$).tw.
6. macrosomia.tw.
7. or/1-6
8. screen$.tw.
9. ((prenatal or early) adj2 diagnosis).tw.
10. blood glucose.tw.
11. risk factor?.tw.
12. (glucose adj (tolerance or intolerance or challenge)).tw.
13. OGTT.tw.
14. GCT.tw.
15. (fasting adj2 glucose).tw.
16. or/8-15
17. specific$.tw.
18. sensitiv$.tw.
19. predictive value.tw.
20. (ROC or "receiver operating characteristic?").tw.
21. accurac$.tw.
22. diagnostic error?.tw.
23. (false adj (negative or positive)).tw.
24. "reproducibility of results".tw.
25. (reference adj2 (standard? or value?)).tw.
26. or/17-25
27. and/7,16,26
28. intervention?.mp.
29. (treating or treatment? or therapy or therapies).mp.
30. manage$.mp.
31. monitor$.mp.
32. sulfonylurea?.tw.
33. gliclazid$.tw.
34. glimepirid$.tw.
35. glipizid$.tw.
36. glyburid$.tw.
37. tolbutamid$.tw.
38. (antidiabet$ or anti-diabet$).tw.
39. insulin?.mp.
40. glibenclamid$.mp.
41. acarbos$.mp.
42. (diet adj2 (therap$ or restrict$ or advice)).tw.
43. medical nutrition$ therapy.tw.
44. MNT.tw.
45. (lifestyle$ or life-style$).mp.
46. (blood glucose adj (self monitor$ or self-monitor$)).tw.
47. ((self monitior$ or self-monitor$) adj blood glucose).tw.
48. SMBG.tw.
49. counsel$.tw.
50. (induc$ adj2 labo?r).tw.
51. c?esarean.tw.
52. pregnanc$ outcome?.tw.
53. or/28-52
54. and/7,53
55. or/27,54

56. clinical trial.pt.
57. randomized controlled trial.pt.
58. randomi?ed.ti,ab.
59. placebo.ti,ab.
60. dt.fs.
61. randomly.ti,ab.
62. trial.ti,ab.
63. groups.ti,ab.
64. or/56-63
65. (animal? not (animal? and human?)).mp.
66. 64 not 65
67. ((cohort? or follow-up or followup or longitud$ or prospectiv$ or retrospectiv$) adj (study or studies or trial?)).tw.
68. 67 not 66
69. (clinical adj2 guideline?).tw.
70. CPG?.tw.
71. ((practice or consensus or position) adj2 (guideline? or recommendation? or statement?)).tw.
72. standard?.tw.
73. protocol?.tw.
74. or/69-73
75. and/55,66 [Clinical trials & RCTs]
76. and/55,68 [Observational studies]
77. and/55,74 [Guidelines]
78. or/75-77
79. limit 78 to (english language and yr="2000-Current")
80. remove duplicates from 79

Table A5. PASCAL

Database: PASCAL via Ovid <1984 to 2011 Week 39>
Search Date: 9 October 2011
Results: 498

1. GDM.tw.
2. (gestation$ adj2 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
3. (pregnan$ adj3 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
4. (maternal adj2 (diabet$ or DM or glyc?emia or hyperglyc?emia)).tw.
5. (hyperglyc?emia adj2 pregnan$).tw.
6. macrosomia.tw.
7. or/1-6
8. screen$.tw.
9. ((prenatal or early) adj2 diagnosis).tw.
10. blood glucose.tw.
11. risk factor?.tw.
12. (glucose adj (tolerance or intolerance or challenge)).tw.
13. OGTT.tw.
14. GCT.tw.
15. (fasting adj2 glucose).tw.
16. or/8-15
17. specific$.tw.
18. sensitiv$.tw.
19. predictive value.tw.
20. (ROC or "receiver operating characteristic?").tw.
21. accurac$.tw.
22. diagnostic error?.tw.
23. (false adj (negative or positive)).tw.
24. "reproducibility of results".tw.
25. (reference adj2 (standard? or value?)).tw.
26. or/17-25
27. and/7,16,26
28. intervention?.mp.
29. (treating or treatment? or therapy or therapies).mp.
30. manage$.mp.
31. monitor$.mp.

32. sulfonylurea?.tw.
33. gliclazid$.tw.
34. glimepirid$.tw.
35. glipizid$.tw.
36. glyburid$.tw.
37. tolbutamid$.tw.
38. (antidiabet$ or anti-diabet$).tw.
39. insulin?.mp.
40. glibenclamid$.mp.
41. acarbos$.mp.
42. (diet adj2 (therap$ or restrict$ or advice)).tw.
43. medical nutrition$ therapy.tw.
44. MNT.tw.
45. (lifestyle$ or life-style$).mp.
46. (blood glucose adj (self monitor$ or self-monitor$)).tw.
47. ((self monitior$ or self-monitor$) adj blood glucose).tw.
48. SMBG.tw.
49. counsel$.tw.
50. (induc$ adj2 labo?r).tw.
51. c?esarean.tw.
52. pregnanc$ outcome?.tw.
53. or/28-52
54. and/7,53
55. or/27,54
56. clinical trial.pt.
57. randomized controlled trial.pt.
58. randomi?ed.ti,ab.
59. placebo.ti,ab.
60. dt.fs.
61. randomly.ti,ab.
62. trial.ti,ab.
63. groups.ti,ab.
64. or/56-63
65. (animal? not (animal? and human?)).mp.
66. 64 not 65
67. ((cohort? or follow-up or followup or longitud$ or prospectiv$ or retrospectiv$) adj (study or studies or trial?)).tw.
68. 67 not 66
69. (clinical adj2 guideline?).tw.
70. CPG?.tw.
71. ((practice or consensus or position) adj2 (guideline? or recommendation? or statement?)).tw.
72. standard?.tw.
73. protocol?.tw.
74. or/69-73
75. and/55,66 [Clinical trials & RCTs]
76. and/55,68 [Observational studies]
77. and/55,74 [Guidelines]
78. or/75-77
79. limit 78 to (english language and yr="2000-Current")
80. remove duplicates from 79

Table A6. Medline In-Process & Other Non-Indexed Citations

Database: Medline In-Process & Other Non-Indexed Citations <October 7, 2011>
Search Date: 7 October 2011
Results: 98

1. GDM.tw.
2. (gestation$ adj2 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
3. (pregnan$ adj3 (diabet$ or DM or glucose intoleran$ or insulin resistan$)).mp.
4. (maternal adj2 (diabet$ or DM or glyc?emia or hyperglyc?emia)).tw.
5. (hyperglyc?emia adj2 pregnan$).tw.
6. macrosomia.tw.
7. or/1-6
8. screen$.tw.

9. ((prenatal or early) adj2 diagnosis).tw.
10. blood glucose.tw.
11. risk factor?.tw.
12. (glucose adj (tolerance or intolerance or challenge)).tw.
13. OGTT.tw.
14. GCT.tw.
15. (fasting adj2 glucose).tw.
16. or/8-15
17. specific$.tw.
18. sensitiv$.tw.
19. predictive value.tw.
20. (ROC or "receiver operating characteristic?").tw.
21. accurac$.tw.
22. diagnostic error?.tw.
23. (false adj (negative or positive)).tw.
24. "reproducibility of results".tw.
25. (reference adj2 (standard? or value?)).tw.
26. or/17-25
27. and/7,16,26
28. intervention?.mp.
29. (treating or treatment? or therapy or therapies).mp.
30. manage$.mp.
31. monitor$.mp.
32. sulfonylurea?.tw.
33. gliclazid$.tw.
34. glimepirid$.tw.
35. glipizid$.tw.
36. glyburid$.tw.
37. tolbutamid$.tw.
38. (antidiabet$ or anti-diabet$).tw.
39. insulin?.mp.
40. glibenclamid$.mp.
41. acarbos$.mp.
42. (diet adj2 (therap$ or restrict$ or advice)).tw.
43. medical nutrition$ therapy.tw.
44. MNT.tw.
45. (lifestyle$ or life-style$).mp.
46. (blood glucose adj (self monitor$ or self-monitor$)).tw.
47. ((self monitior$ or self-monitor$) adj blood glucose).tw.
48. SMBG.tw.
49. counsel$.tw.
50. (induc$ adj2 labo?r).tw.
51. c?esarean.tw.
52. pregnanc$ outcome?.tw.
53. or/28-52
54. and/7,53
55. or/27,54
56. clinical trial.pt.
57. randomized controlled trial.pt.
58. randomi?ed.ti,ab.
59. placebo.ti,ab.
60. dt.fs.
61. randomly.ti,ab.
62. trial.ti,ab.
63. groups.ti,ab.
64. or/56-63
65. (animal? not (animal? and human?)).mp.
66. 64 not 65
67. ((cohort? or follow-up or followup or longitud$ or prospectiv$ or retrospectiv$) adj (study or studies or trial?)).tw.
68. 67 not 66
69. (clinical adj2 guideline?).tw.
70. CPG?.tw.

71. ((practice or consensus or position) adj2 (guideline? or recommendation? or statement?)).tw.
72. standard?.tw.
73. protocol?.tw.
74. or/69-73
75. and/55,66 [Clinical trials & RCTs]
76. and/55,68 [Observational studies]
77. and/55,74 [Guidelines]
78. or/75-77
79. limit 78 to (english language and yr="2000-Current")
80. remove duplicates from 79

A7. CINAHL Plus with Full Text

Database: CINAHL Plus with Full Text via EBSCO <1937–current>
Search Date: 10 October 2011
Results: 275
S39= S35 or S37 or S38
S38= S25 and S33 Limiters - English Language; Published Date from: 20000101-20121231; Exclude MEDLINE records
S37= S25 and S32 Limiters - English Language; Published Date from: 20000101-20121231; Exclude MEDLINE records
S36= S25 and S32
S35= S25 and S31 Limiters - English Language; Published Date from: 20000101-20121231; Exclude MEDLINE records
S34= S25 and S31
S33=(CPG? or "best practice?" or "professional standard?" or "standard of care") OR (practice W2 guideline* or practice W2 recommendation* or practice W2 statement or position W2 guideline* or position W2 recommendation* or position W2 statement or consensus W2 guideline* or consensus W2 recommendation* or consensus W2 statement)
S32=((MH "Prospective Studies+") OR (MH "Retrospective Design")) OR TI (cohort* or follow-up or followup or longitud* or prospectiv* or retrospective*) OR AB (cohort* or follow-up or followup or longitud* or prospectiv* or retrospective*)
S31= S26 or S27 or S28 or S29 or S30
S30=(MH "Placebos") OR TX placebo* OR (MH "Quantitative Studies")
S29= TX randomi* control* trial* OR (MH "Random Assignment") OR TX random* allocat* OR TX allocat* random*
S28= TX clinic* n1 trial* OR (TX ((singl* n1 blind*) or (singl* n1 mask*)) or TX ((doubl* n1 blind*) or (doubl* n1 mask*)) or TX ((tripl* n1 blind*) or (tripl* n1 mask*)) or TX ((trebl* n1 blind*) or (trebl* n1 mask*)))
S27= PT Clinical trial
S26=(MH "Clinical Trials+")
S25= S14 or S24
S24= S5 and S23
S23= S15 or S16 or S17 or S18 or S19 or S20 or S21 or S22
S22=((MH "Labor, Induced") OR (MH "Cesarean Section+") OR (MH "Pregnancy Outcomes")) OR (induc* n2 labo#r or cesarean or caesarean or pregnan* n1 outcome*)
S21=((MH "Counseling") OR (MH "Nutritional Counseling")) OR counsel*
S20=(MH "Blood Glucose Self-Monitoring") OR ("blood glucose" w1 "self monitor*" or "blood glucose" w1 "self-monitor*") OR SMBG
S19=(MH "Life Style Changes") OR (lifestyle* or life-style*)
S18=(MH "Diet Therapy") OR (diet w2 therap* or diet w2 restrict* or diet w2 advice) OR ("medical nutrition therapy" or MNT)
S17=(sulfonyurea? or gliclazid* or glimepirid* or glipizid* or glyburid* or tolbutamid*) OR (antidiabet* or anti-diabet*) OR (insulin* or glibenclamid* or acarbos*)
S16=(MH "Sulfonylurea Compounds+")
S15= intervention* or treating or treatment* or therapy or therapies or manage* or monitor*
S14= S5 and S10 and S13
S13= S11 or S12
S12=(specific* or sensitiv* or predictive w1 value* or accurac* or diagnostic w1 error*) OR (false w1 negative or false w1 positive)
S11=(MH "Diagnostic Errors") OR (MH "Reproducibility of Results") OR (MH "False Negative Results") OR (MH "False Positive Results") OR (MH "Predictive Value of Tests") OR (MH "Sensitivity and Specificity") OR (MH "ROC Curve") OR (MH "Reference Values")
S10= S6 or S7 or S8 or S9
S9=(glucose n1 tolerance or glucose n1 intolerance or glucose n1 challenge) OR (OGTT or GCT) OR fasting w2 glucose
S8=(MH "Glucose Tolerance Test") OR (MH "Blood Glucose Monitoring") OR (MH "Glucose Intolerance") OR (MH "Blood Glucose") OR (MH "Risk Assessment")
S7= screen* OR (prenatal n2 diagnosis or early n2 diagnosis)
S6=(MH "Neonatal Assessment") OR (MH "Health Screening+") OR (MH "Prenatal Diagnosis+")

S5= S1 or S2 or S3 or S4
S4= hyperglyc#emia n2 pregnan* OR macrosomia
S3=(gestation* n2 diabet* or gestation* n2 DM or gestation* n2 glucose intoleran* or gestation* n2 insulin resistan*) OR (pregnan* n2 diabet* or pregnan* n2 DM or pregnan* n2 glucose intoleran* or pregnan* n2 insulin resistan*) OR (maternal n2 diabet* or maternal n2 DM or maternal n2 glucose intoleran* or maternal n2 insulin resistan*)
S2=((MM "Diabetes Mellitus, Gestational") OR (MH "Pregnancy Complications") OR (MH "Fetal Macrosomia")) OR GDM
S1=(MM "Diabetes Mellitus, Gestational") OR (MH "Pregnancy Complications") OR (MH "Fetal Macrosomia")

Table A8. BIOSIS Preview®

Database: Biosis Previews ® via Web of Knowledge[SM] <1926–present>
Search Date: 9 October 2011
Results: 34
17 (#16 OR #15 OR #14) AND Language=(English)
16 (#9) AND Language=(English) AND Document Types=(Meeting OR Meeting Paper) AND Literature Type=(Meeting Abstract OR Meeting Address OR Meeting Paper OR Meeting Poster OR Meeting Report OR Meeting Slide OR Meeting Summary)
15 (#13 AND #9) AND Language=(English
14 (#12 AND #9) AND Language=(English
13 (TS=(CPG* OR "best practice*" OR "professional standard*" OR "standard of care" OR (practice NEAR/2 guideline*) OR (practice NEAR/2 recommendation*) OR (practice NEAR/2 statement) OR (position NEAR/2 guideline*) OR (position NEAR/2 recommendation*) OR (position NEAR/2 statement) OR (consensus NEAR/2 guideline*) OR (consensus NEAR/2 recommendation*) OR (consensus NEAR/2 statement))) AND Language=(English)
12 (#10 NOT #11) AND Language=(English)
11 (TS=(animal* OR rat OR rats OR mouse OR mice OR rodent* OR rabbit OR rabbits OR horse OR horses OR equine OR veterinar* OR bovine OR cow OR cows OR pig OR pigs OR porcine)) AND Language=(English)
10 ((TS=(randomized controlled trial* OR controlled clinical trial* OR research design OR placebo* OR random*) OR TS=(cohort* OR longitude* OR prospectiv* OR retrospectiv* OR long term OR long-term OR longterm OR followup OR "follow up" OR follow-up) AND TS=(study OR studies OR trial))) AND Language=(English)
9 (#8 OR #4) AND Language=(English)
8 (#7 AND #1) AND Language=(English)
7 (#6 OR #5) AND Language=(English)
6 TS= ((diet NEAR/2 therap*) OR (diet NEAR/2 restrict*) OR (diet NEAR/2 advice) OR "medical nutrition* therapy" OR MNT OR lifestyle* OR life-style* OR ("blood glucose" NEAR self-monitor*) OR ("blood glucose" NEAR "self monitor*") OR SMBG OR counsel* OR (induc* NEAR labour) OR (induc* NEAR labor) OR cesarean OR caesarean OR (pregnan* NEAR outcome*))) AND Language=(English)
5 TS= (intervention* OR treat* OR therap* OR sulfonylurea* OR antidiabet* OR anti-diabet* OR gliclazid* OR glimepirid* OR glipizid* OR glyburid* OR tolbutamid* OR antidiabet* OR anti-diabet* OR insulin* OR glibenclamid* OR acarbos*)) AND Language=(English)
4 #3 AND #2 AND #1
3 TS=("sensitivity and specificity" OR sensitiv* OR specific* OR "predictive value" OR (diagnos* NEAR error*) OR "false negative" OR "false positive" OR accurac*)) AND Language=(English)
2 TS=("prenatal screen*" OR (glucose NEAR/3 tolerance) OR (glucose NEAR/3 intoleran*) OR (glucose NEAR/3 challenge*) OR OGTT OR GCT OR "fasting glucose" OR "risk factor* ")) AND Language=(English)
1 TS= ((gestation* NEAR/2 diabet*) OR (gestation* NEAR/2 "glucose intoleran*") OR (gestation* NEAR/2 "insulin resist*") OR (pregnan* NEAR/2 diabet*) OR (pregnan* NEAR/2 "glucose intoleran*") OR (pregnan* NEAR/2 "insulin resist*") OR (maternal NEAR/2 diabet*) OR (maternal NEAR/2 "glucose intoleran*") OR (maternal NEAR/2 "insulin resist*") OR (hyperglycemia NEAR/2 pregnan*) OR (hyperglycaemia NEAR/2 pregnan*) OR macrosomia OR GDM)) AND Language=(English)

Table A9. Science Citation Index Expanded®

Database: Science Citation Index Expanded (SCI-EXPANDED) via Web of Knowledge[SM] <1899–present>
Search Date: 9 October 2011
Results: 2,308
17 (#16 OR #15 OR #14) AND Language=(English)
16 (#9) AND Language=(English) AND Document Types=(Meeting OR Meeting Paper) AND Literature Type=(Meeting Abstract OR Meeting Address OR Meeting Paper OR Meeting Poster OR Meeting Report OR Meeting Slide OR Meeting Summary)
15 (#13 AND #9) AND Language=(English
14 (#12 AND #9) AND Language=(English
13 (TS=(CPG* OR "best practice*" OR "professional standard*" OR "standard of care" OR (practice NEAR/2 guideline*) OR (practice NEAR/2 recommendation*) OR (practice NEAR/2 statement) OR (position NEAR/2 guideline*) OR

(position NEAR/2 recommendation*) OR (position NEAR/2 statement) OR (consensus NEAR/2 guideline*) OR (consensus NEAR/2 recommendation*) OR (consensus NEAR/2 statement))) AND Language=(English)
12 (#10 NOT #11) AND Language=(English)
11 (TS=(animal* OR rat OR rats OR mouse OR mice OR rodent* OR rabbit OR rabbits OR horse OR horses OR equine OR veterinar* OR bovine OR cow OR cows OR pig OR pigs OR porcine)) AND Language=(English)
10 ((TS=(randomized controlled trial* OR controlled clinical trial* OR research design OR placebo* OR random*) OR TS=(cohort* OR longitude* OR prospectiv* OR retrospectiv* OR long term OR long-term OR longterm OR followup OR "follow up" OR follow-up) AND TS=(study OR studies OR trial))) AND Language=(English)
9 (#8 OR #4) AND Language=(English)
8 (#7 AND #1) AND Language=(English)
7 (#6 OR #5) AND Language=(English)
6 TS= ((diet NEAR/2 therap*) OR (diet NEAR/2 restrict*) OR (diet NEAR/2 advice) OR "medical nutrition* therapy" OR MNT OR lifestyle* OR life-style* OR ("blood glucose" NEAR self-monitor*) OR ("blood glucose" NEAR "self monitor*") OR SMBG OR counsel* OR (induc* NEAR labour) OR (induc* NEAR labor) OR cesarean OR caesarean OR (pregnan* NEAR outcome*))) AND Language=(English)
5 TS= (intervention* OR treat* OR therap* OR sulfonylurea* OR antidiabet* OR anti-diabet* OR gliclazid* OR glimepirid* OR glipizid* OR glyburid* OR tolbutamid* OR antidiabet* OR anti-diabet* OR insulin* OR glibenclamid* OR acarbos*)) AND Language=(English)
4 #3 AND #2 AND #1
3 TS=("sensitivity and specificity" OR sensitiv* OR specific* OR "predictive value" OR (diagnos* NEAR error*) OR "false negative" OR "false positive" OR accurac*)) AND Language=(English)
2 TS=("prenatal screen*" OR (glucose NEAR/3 tolerance) OR (glucose NEAR/3 intoleran*) OR (glucose NEAR/3 challenge*) OR OGTT OR GCT OR "fasting glucose" OR "risk factor* ")) AND Language=(English)
1 TS= ((gestation* NEAR/2 diabet*) OR (gestation* NEAR/2 "glucose intoleran*") OR (gestation* NEAR/2 "insulin resist*") OR (pregnan* NEAR/2 diabet*) OR (pregnan* NEAR/2 "glucose intoleran*") OR (pregnan* NEAR/2 "insulin resist*") OR (maternal NEAR/2 diabet*) OR (maternal NEAR/2 "glucose intoleran*") OR (maternal NEAR/2 "insulin resist*") OR (hyperglycemia NEAR/2 pregnan*) OR (hyperglycaemia NEAR/2 pregnan*) OR macrosomia OR GDM)) AND Language=(English)

Table A10. Conference Proceedings Citation Index–Science

Database: Conference Proceedings Citation Index- Science [CPCI-S] via Web of Science^{SM} <1990–present>
Search Date: 9 October 2011
Results: 562
17 (#16 OR #15 OR #14) AND Language=(English)
16 (#9) AND Language=(English) AND Document Types=(Meeting OR Meeting Paper) AND Literature Type=(Meeting Abstract OR Meeting Address OR Meeting Paper OR Meeting Poster OR Meeting Report OR Meeting Slide OR Meeting Summary)
15 (#13 AND #9) AND Language=(English
14 (#12 AND #9) AND Language=(English
13 (TS=(CPG* OR "best practice*" OR "professional standard*" OR "standard of care" OR (practice NEAR/2 guideline*) OR (practice NEAR/2 recommendation*) OR (practice NEAR/2 statement) OR (position NEAR/2 guideline*) OR (position NEAR/2 recommendation*) OR (position NEAR/2 statement) OR (consensus NEAR/2 guideline*) OR (consensus NEAR/2 recommendation*) OR (consensus NEAR/2 statement))) AND Language=(English)
12 (#10 NOT #11) AND Language=(English)
11 (TS=(animal* OR rat OR rats OR mouse OR mice OR rodent* OR rabbit OR rabbits OR horse OR horses OR equine OR veterinar* OR bovine OR cow OR cows OR pig OR pigs OR porcine)) AND Language=(English)
10 ((TS=(randomized controlled trial* OR controlled clinical trial* OR research design OR placebo* OR random*) OR TS=(cohort* OR longitude* OR prospectiv* OR retrospectiv* OR long term OR long-term OR longterm OR followup OR "follow up" OR follow-up) AND TS=(study OR studies OR trial))) AND Language=(English)
9 (#8 OR #4) AND Language=(English)
8 (#7 AND #1) AND Language=(English)
7 (#6 OR #5) AND Language=(English)
6 TS= ((diet NEAR/2 therap*) OR (diet NEAR/2 restrict*) OR (diet NEAR/2 advice) OR "medical nutrition* therapy" OR MNT OR lifestyle* OR life-style* OR ("blood glucose" NEAR self-monitor*) OR ("blood glucose" NEAR "self monitor*") OR SMBG OR counsel* OR (induc* NEAR labour) OR (induc* NEAR labor) OR cesarean OR caesarean OR (pregnan* NEAR outcome*))) AND Language=(English)
5 TS= (intervention* OR treat* OR therap* OR sulfonylurea* OR antidiabet* OR anti-diabet* OR gliclazid* OR glimepirid* OR glipizid* OR glyburid* OR tolbutamid* OR antidiabet* OR anti-diabet* OR insulin* OR glibenclamid* OR acarbos*)) AND Language=(English)
4 #3 AND #2 AND #1
3 TS=("sensitivity and specificity" OR sensitiv* OR specific* OR "predictive value" OR (diagnos* NEAR error*) OR "false negative" OR "false positive" OR accurac*)) AND Language=(English)

2 TS=("prenatal screen*" OR (glucose NEAR/3 tolerance) OR (glucose NEAR/3 intoleran*) OR (glucose NEAR/3 challenge*) OR OGTT OR GCT OR "fasting glucose" OR "risk factor* ")) AND Language=(English)
1 TS= ((gestation* NEAR/2 diabet*) OR (gestation* NEAR/2 "glucose intoleran*") OR (gestation* NEAR/2 "insulin resist*") OR (pregnan* NEAR/2 diabet*) OR (pregnan* NEAR/2 "glucose intoleran*") OR (pregnan* NEAR/2 "insulin resist*") OR (maternal NEAR/2 diabet*) OR (maternal NEAR/2 "glucose intoleran*") OR (maternal NEAR/2 "insulin resist*") OR (hyperglycemia NEAR/2 pregnan*) OR (hyperglycaemia NEAR/2 pregnan*) OR macrosomia OR GDM)) AND Language=(English)

Table A11. LILACS (Latin American and Caribbean Health Science Literature)

Database: LILACS (Latin American and Caribbean Health Science Literature) <1982–current>
Search Date: 14 October 2011
Results: 236
1. gestational diabet$ AND (screening OR diagnos$)
2. maternal diabet$ AND (screening OR diagnos$)
3. gestational diabet$ AND (treating or treatment$ or therapy or therapies)
4. maternal diabet$ AND (treating or treatment$ or therapy or therapies)

Table A12. OCLC PapersFirst and PapersFirst

Databases:
ProceedingsFirst
PapersFirst
Search Date: 16 October 2011
Results:
ProceedingsFirst: 138; PapersFirst: 102
(kw: gestation* w2 diabet* OR kw: gestation* w2 glucose w intoleran* OR kw: gestation* w2 insulin w resist* OR kw: pregnan* w2 diabet* OR kw: pregnan* w2 glucose w intoleran* OR kw: pregnan* w2 insulin w resist* OR kw: maternal w2 diabet* OR kw: maternal w2 glucose w intoleran* OR kw: maternal w2 insulin w resist* OR kw: hyperglycemia w2 pregnan* OR kw: hyperglycaemia w2 pregnan* OR kw: macrosomia OR kw: GDM) and ((kw: prenatal w screen* OR kw: glucose w3 tolerance OR kw: glucose w3 intoleran* OR kw: glucose w3 challenge* OR kw: OGTT OR kw: GCT OR kw: fasting w glucose OR kw: risk w factor*) or ((kw: intervention* OR kw: treat* OR kw: therap* OR kw: sulfonylurea* OR kw: antidiabet* OR kw: anti-diabet* OR kw: gliclazid* OR kw: glimepirid* OR kw: glipizid* OR kw: glyburid* OR kw: tolbutamid* OR kw: antidiabet* OR kw: anti-diabet* OR kw: insulin* OR kw: glibenclamid* OR kw: acarbos*) or (kw: diet w2 therap* OR kw: diet w2 restrict* OR kw: diet w2 advice OR kw: medical w nutrition* w therapy OR kw: MNT OR kw: lifestyle* OR kw: life-style* OR kw: blood w glucose w self-monitor* OR kw: blood w glucose w self w monitor* OR kw: SMBG OR kw: counsel* OR kw: induc* w labour OR kw: induc* w labor OR kw: cesarean OR kw: caesarean OR kw: pregnan* w outcome*)))

Table A13. PubMed

Database: PubMed via NLM <last 180 days from 9 October 2011>
Search Date: 9 October 2011
Results: 377
#46 #39 NOT #45
#45 animal[TI] OR rat[TI] OR rats[TI] OR mouse [TI] OR mice[TI] OR rodent*[TI] OR rabbit*[TI] OR horse*[TI] OR horses[TI] veterinar*[TI] OR cattle[TI] OR bovine[TI] OR cow[TI] OR cows[TI] OR swine[TI] OR pig[TI] OR pigs[TI] OR porcine[TI]
#39 #21 OR #37 Limits: English, published in the last 180 days
#38 #21 OR #37
#37 #7 and #36
#36 #22 OR #23 OR #25 OR #28 OR #30 OR #31 OR #32 OR #33 OR #34 OR #34
#35 pregnanc* outcome*
#34 cesarean OR caesarean
#33 ((induc* AND labour) OR (induc* AND labor))
#32 counsel*
#31 SMBG
#30 ((self monitor* OR self-monitor*) AND blood glucose)
#28 (blood glucose AND (self monitor* OR self-monitor*))
#25 lifestyle OR life-style
#24 diet therap* OR diet* restrict* OR diet* advice OR medical nutrition therapy OR MNT
#23 sulfonylurea* OR gliclazid* OR glimepirid* OR glipizid* OR glyburid* OR tolbutamid* OR antidiabet* OR anti-diabet* OR insulin* OR glibenclamid* OR acarbos*
#22 intervention* OR treating OR treatment? OR therapy OR therapies OR manage* OR monitor*

#21 #7 AND #16 AND #20
#20 #17 OR #18 OR #19
#19 reference standard* OR reference value*
#18 ROC OR "receiver operating characteristic"
#17 specific* OR sensitiv* OR predictive value OR accurac* OR diagnostic error*
#16 #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15
#15 fasting glucose
#14 OGTT OR GCT
#13 (glucose AND (tolerance OR intolerance OR challenge))
#12 risk factor*
#11 blood glucose
#10 ((prenatal OR early) AND diagnosis)
#9 screen*
#8 mass screening[MeSH Terms]
#7 #1 OR #2 OR #3 OR #4 OR #5 OR #6
#6 macrosomia
#5 ((hyperglycaemia OR hyperglycemia) AND pregnan*)
#4 (maternal AND (diabetic* OR diabete* OR DM OR glucose intoleran* OR insulin resistan*))
#3 (pregnan* AND (diabetic* OR diabete* OR DM OR glucose intoleran* OR insulin resistan*))
#2 (gestation* AND (diabetic* OR diabete* OR DM OR glucose intoleran* OR insulin resistan*))
#1 GDM

Table A14. Clinical Trials.gov and WHO

Databases:
ClinicalTrials.gov <1987 to February week 3 2012>
WHO International Clinical Trials Registry
Search Date: 23 February 2012
Results: 200
((asperger) OR (autistic disorder) OR autism OR schizophrenia OR (bipolar disorder) OR (depression) OR (bipolar disorder) OR (obsessive-compulsive) OR (post-traumatic) OR (anorexia nervosa) OR anorexia) AND (antipsychotics) AND (child OR adolescent OR pediatric OR infant) AND PDN(>1/1/1987) AND PDN(<12/31/2010)

Appendix B. Review Forms

B1. Screening Criteria for Key Questions 1-5
B2. Eligibility Criteria for Key Questions 1-5
B3. Methodological Quality Assessment by Study Design
    a. Diagnostic studies – QUADAS-2 Tool
    b. Randomized controlled trials – Cochrane Collaboration's tool for assessing risk of bias
    c. Cohort studies – Newcastle-Ottawa Quality Assessment Scale
B4. Data Extraction Forms
    a. Screening and diagnosing gestational diabetes – key question 1
    b. Screening and diagnosing gestational diabetes – key question 2
    c. Screening and diagnosing gestational diabetes – key question 3
    d. Screening and diagnosing gestational diabetes – key question 4 and 5

B1. Screening Criteria for Key Questions 1-5

1. Primary Research	*Yes*	*No*	*Unclear*
2. Published in English language	*Yes*	*No*	*Unclear*
3. Published from 1995 onward	*Yes*	*No*	*Unclear*
4. Must have a comparison group (i.e., RCT, NRCT, R or P cohort, case control)	*Yes*	*No*	*Unclear*
5. ***Population****:* Pregnant women	*Yes*	*No*	*Unclear*
6. ***Intervention***: Using **any** GDM screening or diagnostic approach, (e.g., 1-step, 2-step, or other); ***and/or*** Any treatment for GDM (e.g., dietary advice, blood glucose monitoring, insulin therapy)	*Yes*	*No*	*Unclear*

Notes for screeners:

1. Mark each study as "no" [exclude], "unclear" or "yes" [retrieve full text] based on the criteria above.
2. FLAG any relevant systematic reviews or meta-analyses using the code "sr".
3. FLAG any studies that may be useful for background information with the code "bkg".

Key words have been colour-coded and will appear in a different font. Here is an index of the colouring:

Green → population (e.g., gestational diabetes, pregnancy)
Purple → treatments (e.g., diet, insulin, blood glucose monitoring, antidiabetic)
Aqua → screening-related terms (e.g., screening, diagnosis, sensitivity, specificity, positive predictive value, negative predictive value)
Orange → specific tests (e.g., glucose tolerance test, glucose challenge test, glucose screening test, diagnostic threshold)
Blue → study designs (e.g., randomized, controlled trial, cohort, case control)

B2. Eligibility Criteria for Key Questions 1-5

INCLUSION / EXCLUSION FORM

Reviewer:	**Ref ID:**		
CRITERIA	**Yes**	**No**	**Unclear**
1. PUBLICATION TYPE			
a) Report of primary research	☐	☐	☐
b) Full report available (Exclude abstracts and conference proceedings)	☐	☐	☐
c) English language	☐	☐	☐
d) Published in 1995 onward	☐	☐	☐
2. STUDY DESIGN			
a) Comparative study design (2 or more groups); one of: i. RCT ii. NRCT iii. Prospective or retrospective cohort studies (with concurrent or nonconcurrent/historical control groups)	☐	☐	☐
3. POPULATION			
a) Pregnant women (any duration of gestation); *Exclude* if >20% of enrolled women had known pre-existing diabetes **and** no subgroup analysis	☐	☐	☐
4. INTERVENTION			
a) Evaluating **any** GDM screening or diagnostic approach, (KQ1 & 2) or screening / diagnostic threshold (KQ3) *and/or* b) Evaluating **any** treatment for GDM (KQ4 & 5)	☐	☐	☐
5. COMPARATORS			
One or more of the following: a) Any reference standard, other screening / diagnostic test, or criteria (KQ1) [note: can also be a risk-factor if used for screening]; b) No screening / diagnostic test for GDM (KQ2); c) Patients meeting different screening / diagnostic threshold for GDM (e.g., GDM vs. no GDM) (KQ3); d) Placebo or no treatment (KQ4 & 5) *Exclude studies that compare 2 or more treatment, but have no placebo, standard care or no treatment group*	☐	☐	☐
6. OUTCOME			
Any one or more of the following: a) Test properties (i.e., sensitivity, specificity, predictive values, accuracy; *not* yield only); b) Maternal outcomes: i. Short-term: preeclampsia/maternal hypertension, cesarean delivery, depression, birth trauma, mortality, weight gain, other morbidity ii. Long-term: Type 2 DM risk, obesity, hypertension c) Fetal/neonatal/child outcomes: i. Short-term: macrosomia, shoulder dystocia, clavicular fracture, brachial plexus injury, birth injury, hypoglycemia, hyperbilirubinemia, mortality, other morbidity ii. Long-term: obesity, type 2 DM, transgenerational GDM d) Any adverse events or harms of screening or treatment (e.g., anxiety, healthcare system issues, burden on practitioner's office, increased interventions, postpartum depression, small for gestational age, costs, resource allocations)	☐	☐	☐
Comments:			
REVIEWER'S DECISION : Include ☐ Exclude ☐ Unsure ☐			
RELEVANT TO QUESTION(S): KQ1 ☐ KQ2 ☐ KQ3 ☐ KQ4 ☐ KQ5 ☐			

B3. Methodological Quality

a. QUADAS-2 Checklist (Diagnostic Studies)

Item	Assessment
1. *Patient Selection*	
a. Was a consecutive or random sample of patients enrolled? *Support for judgment*	
b. Did the study avoid inappropriate exclusions? *Support for judgment*	
c. Was the study a low risk of bias?	
d. Is the study applicable to the review? *Support for judgment*	
2. *Index Test*	
a. Were the index test results interpreted without knowledge of the results of the reference standard? *Support for judgment*	
b. If a threshold was used, was it pre-specified? *Support for judgment*	
c. Was the study a low risk of bias?	
d. Is the study applicable to the review? *Support for judgment*	
3. *Reference Standard*	
a. Is the reference standard likely to correctly classify the target audience? *Support for judgment*	
b. Were the reference standard results interpreted without knowledge of the results of the index test? *Support for judgment*	
c. Was the study a low risk of bias?	
d. Is the study applicable to the review? *Support for judgment*	
4. *Flow and Timing*	
a. Was there an appropriate interval between the index test and reference standard? *Support for judgment*	
b. Did all patients receive the same reference standard? *Support for judgment*	
c. Were all patients included in the analysis? *Support for judgment*	
d. Was the study a low risk of bias?	

b. The Cochrane Collaboration's tool for assessing risk of bias (randomized controlled trials)

Domain	Description	Review authors' judgment
Random sequence generation		Was the allocation sequence adequately generated?
Allocation concealment		Was allocation adequately concealed?
Blinding of participants and personnel	Subjective outcomes	Was knowledge of the allocated intervention adequately prevented during the study? Subjective: Objective:
	Objective outcomes	
Blinding of outcome assessment	Subjective outcomes	Was knowledge of the allocated intervention adequately prevented during the study? Subjective: Objective:
	Objective outcomes	
Incomplete outcome data, Outcome:	Subjective outcomes	Were incomplete outcome data adequately addressed? Subjective: Objective:
	Objective outcomes	
Selective outcome reporting		Are reports of the study free of suggestion of selective outcome reporting?
Other sources of bias		Was the study apparently free of other problems that could put it at a high risk of bias?
Overall risk of bias	Subjective outcomes	
	Objective outcomes	

c. Newcastle-Ottawa Quality Assessment Scale (Cohort Studies)

Selection

1) Representativeness of the exposed cohort (i.e., glucose intolerant or GDM patients)
   a) truly representative of the average patient with glucose intolerance in the community *****
   b) somewhat representative of the average glucose intolerance in the community *****
   c) selected group of users eg nurses, volunteers
   d) no description of the derivation of the cohort
2) Selection of the non-exposed cohort (i.e., normal or minimal glucose intolerant patients)
   a) drawn from the same community as the exposed cohort *****
   b) drawn from a different source
   c) no description of the derivation of the non exposed cohort
3) Ascertainment of exposure
   a) secure record (eg surgical records) *****
   b) structured interview *****
   c) written self report
   d) no description
4) Demonstration that outcome of interest was not present at start of study
   a) yes *****
   b) no

Comparability

1) Comparability of cohorts on the basis of the design or analysis
   a) study controls for age, race/ethnicity, weight/BMI, previous GDM, or family history of diabetes ******
   b) study controls for any additional factor *****

Outcome

1) Assessment of outcome
   a) independent blind assessment *****
   b) record linkage *****
   c) self report
   d) no description
2) Was follow-up long enough for outcomes to occur
   a) yes (follows patients at least until birth) *****
   b) no
3) Adequacy of follow up of cohorts
   a) complete follow up - all subjects accounted for *****
   b) subjects lost to follow up unlikely to introduce bias: small number lost (>90% follow up), or description provided of those lost *****
   c) follow up rate <75% and no description of those lost
   d) no statement

TOTAL STARS (0-9)

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

B4. Data Extraction

a. Screening and Diagnosing Gestational Diabetes – Key Question 1

I. Coder Information

RefID:	First Author:	Year:
DE initials:	DV initials:	Other KQs: ☐2; ☐3; ☐4; ☐5

II. Study Characteristics

Country:	Publication type:	Study design:
Centers:	Recruitment start date (e.g., Jan 1998):	Recruitment end date (e.g., Feb 2000):
Funding: ☐Industry; ☐Government; ☐Academic; ☐Foundation; ☐No funding; ☐Other; ☐ND		
If industry, specify firm*:	If "other," specify*:	
Blinding to test result:	Duration of followup:	

* Use "NR" if not reported

III. Selection Criteria and Testing Conditions

Inclusion criteria:	Exclusion criteria: Exclude pre-pregnancy (type 1, 2)? Exclude overt diabetes diagnosed during pregnancy?
Did patients routinely undergo early testing for overt diabetes during pregnancy?	
Patients Enrolled Consecutively: Yes ☐ No ☐ ND ☐	Comparisons Done: Matched Study (all comparator tests performed in all patients) ☐ Random (comparator tests done in different patients) ☐ Non-Random (comparator tests in different patients, select gp) ☐
Reference standard reported?	If so, specify:

IV. Screening and Diagnostic Tests

GCT/GST?	**OGTT?**	**Other test 1?**	**Other test 2?**
Index test? : Pre-test protocol (fast/diet): Test Intervals: ☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr Glucose load: Time of test (wks): Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR Brand of beverage*: Amount of liquid*:	Index test? : Pre-test protocol (fast/diet): Test Intervals: ☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr Glucose load: Time of test (wks): Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR Brand of beverage*: Amount of liquid*:	Specify: Index test? : Pre-test protocol (fast/diet): Test Intervals: ☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr Glucose load: Time of test (wks): Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR Brand of beverage*: Amount of liquid*:	Specify: Index test? : Pre-test protocol (fast/diet): Test Intervals: ☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr Glucose load: Time of test (wks): Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR Brand of beverage*: Amount of liquid*:
Brand of Glucose meter: Manufacturing company:		Measurements performed by trained staff?	
Plasma glucose estimation method: Manufacturing company:			
Central lab?	*Notes:*		

***If not reported, use NR**

V. Study Arms

	Group 1	Group 2	Group 3	Group 4	TOTAL
Group label					
GCT: Fasting	±	±	±	±	±
GCT: 1hr	±	±	±	±	±
GCT: 2hr	±	±	±	±	±
GCT:3hr	±	±	±	±	±
OGTT: Fasting	±	±	±	±	±
OGTT: 1hr	±	±	±	±	±
OGTT: 2hr	±	±	±	±	±
OGTT: 3hr	±	±	±	±	±
Treatment status					
Glucose levels reported in the following units: ☐mg/dL; ☐mmol/L			Glucose levels reported as: ☐mean ± SD; ☐median ± IQR		
Are groups mutually exclusive?					

I. Baseline Characteristics

	Group 1	Group 2	Group 3	Group 4	TOTAL
Pts enrolled, n					
Pts analyzed, n					
Withdrawals, n					
Age (yr), ☐mean ± SD ☐median ± IQR	±	±	±	±	±
Prepregn. weight, ☐lb; ☐kg ☐mean ± SD ☐median ± IQR	±	±	±	±	±
BMI, ☐ mean±SD ☐median ± IQR	±	±	±	±	±
SBP (mmHg), ☐mean ± SD ☐median ± IQR	±	±	±	±	±
	Group 1	**Group 2**	**Group 3**	**Group 4**	**TOTAL**
White, n					
Black, n					
Hispanic, n					
Asian, n					
Other, n					
Gestation at time of test (wk) ☐mean ± SD ☐median ± IQR	±	±	±	±	±
Smoking, n					
Alcohol use, n					
Family history of					

diabetes, n					
History of GDM, n					
Parity, n	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2
Parity ☐mean ± SD ☐median ± IQR	±	±	±	±	±
Comorbidities, n					
Comments					

II. Conclusions

Briefly paraphrase the author conclusions:

REFERENCES TO BE CHECKED:

ASSOCIATED PUBLICATIONS (list all separated by semi-colons):

b. Screening and Diagnosing Gestational Diabetes – Key Question 2

IV. Coder Information

RefID:	First Author:	Year:
DE initials:	DV initials:	Other KQs: ☐1; ☐3; ☐4; ☐5

V. Study Characteristics

Country:	Publication type:	Study design:
Centers:	Recruitment start date (e.g., Jan 1998):	Recruitment end date (e.g., Feb 2000):
Funding: ☐Industry; ☐Government; ☐Academic; ☐Foundation; ☐No funding; ☐Other; ☐ND		
If industry, specify firm*:	If "other," specify*:	
Blinding to test result:	Duration of followup:	

* Use "NR" if not reported

VI. Study Eligibility Criteria

Inclusion criteria:	Exclusion criteria: Exclude pre-pregnancy (type 1, 2)? Exclude overt diabetes diagnosed during pregnancy?
Did patients routinely undergo early testing for overt diabetes during pregnancy?	

VII. Screening and Diagnostic Tests

GCT/GST?	**OGTT?**	**Other test?** Specify:
Test intervals:	Test intervals:	Test intervals:
☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr	☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr	☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr
Glucose load:	Glucose load:	Glucose load:
Time of test (wks):	Time of test (wks):	Time of test (wks):
Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR	Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR	Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR
Central lab?	*Notes:*	

VIII. Study Arms

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	TOTAL
Group label								
GCT: Fasting	±	±	±	±	±	±	±	±
GCT: 1 hr	±	±	±	±	±	±	±	±
GCT: 2 hr	±	±	±	±	±	±	±	±
GCT: 3 hr	±	±	±	±	±	±	±	±
OGTT: Fasting	±	±	±	±	±	±	±	±
OGTT: 1 hr	±	±	±	±	±	±	±	±
OGTT: 2 hr	±	±	±	±	±	±	±	±
OGTT: 3 hr	±	±	±	±	±	±	±	±
Treatment status								
Glucose levels reported in the following units: ☐mg/dL; ☐mmol/L				Glucose levels reported as: ☐mean ± SD; ☐median ± IQR				
Are groups mutually exclusive?								

IX. Baseline Characteristics

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	TOTAL
Pts enrolled, n								
Pts analyzed, n								
Withdrawals, n								
Age (yr), ☐mean ± SD ☐median ± IQR	±	±	±	±	±	±	±	±
Prepregn. weight, ☐lb; ☐kg ☐mean ± SD ☐median ± IQR	±	±	±	±	±	±	±	±
BMI, ☐ mean±SD ☐median ± IQR	±	±	±	±	±	±	±	±
SBP (mmHg), ☐mean ± SD ☐median ± IQR	±	±	±	±	±	±	±	±
	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	TOTAL

White, n								
Black, n								
Hispanic, n								
Asian, n								
Other, n								
Gestation at time of test (wk) ☐mean ± SD ☐median ± IQR	±	±	±	±	±	±	±	±
Smoking, n								
Alcohol use, n								
Family history of diabetes, n								
History of GDM, n								
Parity, n	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2
Parity ☐mean ± SD ☐median ± IQR	±	±	±	±	±	±	±	±
Comorbidities, n								
Comments								

X. Conclusions

Briefly paraphrase the author conclusions:

REFERENCES TO BE CHECKED:

ASSOCIATED PUBLICATIONS (list all separated by semi-colons):

c. Screening and Diagnosing Gestational Diabetes – Key Question 3

I. Coder Information

RefID:	First Author:	Year:
DE initials:	DV initials:	Other KQs: ☐1; ☐2; ☐4; ☐5

II. Study Characteristics

Country:	Publication type:	Study design:
Centers:	Recruitment start date (e.g., Jan 1998):	Recruitment end date (e.g., Feb 2000):
Funding: ☐Industry; ☐Government; ☐Academic; ☐Foundation; ☐No funding; ☐Other; ☐ND		
If industry, specify firm*:	If "other," specify*:	
Blinding to test result:	Duration of followup:	

* Use "NR" if not reported

III. Study Eligibility Criteria

Inclusion criteria:	Exclusion criteria: Exclude pre-pregnancy (type 1, 2)? Exclude overt diabetes diagnosed during pregnancy?
Did patients routinely undergo early testing for overt diabetes during pregnancy?	

IV. Screening and Diagnostic Tests

GCT/GST?	**OGTT?**	**Other test?**	Specify:
Test intervals:	Test intervals:	Test intervals:	
☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr	☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr	☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr	
Glucose load:	Glucose load:	Glucose load:	
Time of test (wks):	Time of test (wks):	Time of test (wks):	
Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR	Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR	Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR	
Central lab?	*Notes:*		

V. Study Arms

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	TOTAL
Group label								
GCT: Fasting	±	±	±	±	±	±	±	±
GCT: 1 hr	±	±	±	±	±	±	±	±
GCT: 2 hr	±	±	±	±	±	±	±	±
GCT: 3 hr	±	±	±	±	±	±	±	±
OGTT: Fasting	±	±	±	±	±	±	±	±
OGTT: 1 hr	±	±	±	±	±	±	±	±
OGTT: 2 hr	±	±	±	±	±	±	±	±
OGTT: 3 hr	±	±	±	±	±	±	±	±
Treatment status								
Glucose levels reported in the following units: ☐mg/dL; ☐mmol/L				Glucose levels reported as: ☐mean ± SD; ☐median ± IQR				
Are groups mutually exclusive?								

VI. Baseline Characteristics

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	TOTAL
Pts enrolled, n								
Pts analyzed, n								
Withdrawals, n								
Age (yr), ☐mean ± SD ☐median ± IQR	±	±	±	±	±	±	±	±
Prepregn. weight, ☐lb; ☐kg ☐mean ± SD ☐median ± IQR	±	±	±	±	±	±	±	±
BMI, ☐ mean±SD ☐median ± IQR	±	±	±	±	±	±	±	±
SBP (mmHg), ☐mean ± SD ☐median ± IQR	±	±	±	±	±	±	±	±
	Group 1	**Group 2**	**Group 3**	**Group 4**	**Group 5**	**Group 6**	**Group 7**	**TOTAL**

White, n								
Black, n								
Hispanic, n								
Asian, n								
Other, n								
Gestation at time of test (wk) ☐mean ± SD ☐median ± IQR	±	±	±	±	±	±	±	±
Smoking, n								
Alcohol use, n								
Family history of diabetes, n								
History of GDM, n								
Parity, n	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2
Parity ☐mean ± SD ☐median ± IQR	±	±	±	±	±	±	±	±
Comorbidities, n								
Comments								

VII. Conclusions

Briefly paraphrase the author conclusions:

REFERENCES TO BE CHECKED:

ASSOCIATED PUBLICATIONS (list all separated by semi-colons):

d. Screening and Diagnosing Gestational Diabetes Mellitus – Key Question 4 and 5

I. Coder Information

Ref ID:	First Author:	Year of Publication:
DE Initials:	DE Reviewer Initials:	Other KQs: ☐1; ☐2; ☐3

II. Study Characteristics

Country:	Publication Type:	Study Design:
Centers:	Recruitment start date (e.g. Jan. 2001):	Recruitment end date (e.g. Feb. 2000):
Funding: ☐Industry; ☐Government; ☐Academic; ☐Foundation; ☐No funding; ☐Other; ☐ND		
If Industry, specify firm:*	If "Other", specify*:	
Blinding to test result:	Duration of followup:	

*use NR if not reported

III. Study Eligibility Criteria

Inclusion Criteria:	Exclusion Criteria: Exclude pre-pregnancy diabetes (type 1, 2)? Exclude overt diabetes diagnosed during pregnancy?
Did patients routinely undergo early testing for overt diabetes during pregnancy?	

IV. Screening and Diagnostic Tests

GCT/GST?	**OGTT?**	**Other test?**	Specify:
Test intervals:	Test intervals:	Test intervals:	
☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr	☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr	☐Fasting; ☐1 hr; ☐2 hr; ☐ 3 hr	
Glucose load:	Glucose load:	Glucose load:	
Time of test (wks):	Time of test (wks):	Time of test (wks):	
Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR	Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR	Criteria: ☐ ADA, year: ☐ CC, year: ☐ NDDG, year: ☐ WHO, year: ☐ Other1: , year: ☐ Other2: , year: ☐ NR	
Central lab?	*Notes:*		

V. Study Arms

	Group 1	Group 2	Group 3	Group 4	TOTAL
Group label					
GCT: Fasting	±	±	±	±	±
GCT: 1hr	±	±	±	±	±
GCT: 2hr	±	±	±	±	±
GCT:3hr	±	±	±	±	±
OGTT: Fasting	±	±	±	±	±
OGTT: 1hr	±	±	±	±	±
OGTT: 2hr	±	±	±	±	±
OGTT: 3hr	±	±	±	±	±
Treatment status					
Glucose levels reported in the following units: ☐mg/dL; ☐mmol/L			Glucose levels reported as: ☐mean ± SD; ☐median ± IQR		
Are groups mutually exclusive?					

VI. Intervention

	Group 1	Group 2	Group 3	Group 4
Study arm label				
Brief description of intervention				
Care provider(s)				
BG target: FGB	Units:	Units:	Units:	Units:
BG target: 1 hr	Units:	Units:	Units:	Units:
Dietary counseling/ advice?				
Formal diet plan?				
If formal diet, describe:				
Involve dietician/ nutritionist?				
BG monitoring?				
Frequency of BG monitoring	x per	x per	x per	x per
BGM device				
Insulin?				
Oral medications?				
Drug name				
BG values for prescription:	≥ Units: Time:	≥ Units: Time:	≥ Units: Time:	≥ Units: Time:
Dosing				
Daily dosage	Units:	Units:	Units:	Units:
Other tx				
Rules for tx/dose adjustment				

Comments				

VII. Baseline Characteristics

	Group 1	**Group 2**	**Group 3**	**Group 4**	**TOTAL**
Pts enrolled, n					
Pts analyzed, n					
Withdrawals, n					
Age (yr) ☐mean ± SD ☐median ± IQR	±	±	±	±	±
Prepregn. weight, ☐lb; ☐kg ☐mean ± SD ☐median ± IQR	±	±	±	±	±
BMI, ☐ mean±SD ☐median ± IQR	±	±	±	±	±
SBP (mmHg), ☐ mean ± SD ☐ median ± IQR	±	±	±	±	±
White, n					
Black, n					
Hispanic, n					
Asian, n					
Other, n					
Gestation at time of test (wk) ☐mean ± SD ☐median ± IQR	±	±	±	±	±
Smoking, n					
Alcohol use, n					
Family history of diabetes, n					
History of GDM, n					
Parity, n	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2	0 1 ≥2
Parity ☐ mean ± SD ☐ median ± IQR	±	±	±	±	±
Overt diabetes, n					
Comorbidities, n					

Comments					

VIII. Conclusions

Briefly paraphrase author conclusions:

REFERENCES TO BE CHECKED:

ASSOCIATED PUBLICATIONS (list all separated by semi-colons):

Appendix C. Methodological Quality of Included Studies

Table C1. Methodological quality of diagnostic studies using QUADAS-2 for Key Question 1

Author, Year Study design	1. Patient Selection*				2. Index Test*				3. Reference Standard*				4. Flow and Timing*			
	a. sample	b. exclusion	c. low risk of bias	d. applicable	a. reference results not known	b. threshold	c. low risk of bias	d. applicable	a. likely to classify	b. index results not known	c. low risk of bias	d. applicable	a. interval	b. same standard	c. all included in analysis	d. low risk of bias
Agarwal, 2000 PCS	No	Yes	No	No	U	U	U	Yes	Yes	U	U	Yes	U	Yes	Yes	U
Agarwal, 2001 PCS (34426)	No	Yes	No	No	U	Yes	U	Yes	Yes	U	U	Yes	U	Yes	Yes	U
Agarwal, 2005a PCS	Yes	U	U	No	U	Yes	U	Yes	Yes	U	U	Yes	U	Yes	Yes	U
Agarwal, 2005b PCS	Yes	Yes	Yes	No	U	Yes	U	Yes	Yes	U	U	Yes	U	Yes	Yes	U
Agarwal, 2006 PCS	Yes	Yes	Yes	No	U	U	U	Yes	Yes	U	U	Yes	U	Yes	Yes	U
Agarwal, 2008 PCS	Yes	No	U	No	U	Yes	U	Yes	Yes	U	U	Yes	Yes	Yes	Yes	Yes
Agarwal, 2011 PCS	Yes	Yes	Yes	No	U	U	U	U	Yes	U	U	Yes	Yes	Yes	Yes	Yes
Ardawi, 2000 PCS	Yes	U	U	No	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	U	U
Ayach, 2006 PCS	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	U	U	Yes	Yes	Yes	U	U
Balaji(1), 2011 PCS	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Balaji(2), 2011 PCS	Yes	Yes	Yes	No	U	Yes	U	Yes	U	Yes	U	Yes	Yes	Yes	U	U
Berkus, 1995 PCS	No	Yes	No	Yes	U	Yes	U	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bobrowski, 1996 PCS	No	Yes	No	Yes	U	Yes	U	Yes	Yes	U	U	Yes	U	Yes	Yes	U
Brustman, 1995 PCS	No	Yes	No	Yes	U	Yes	U	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Author, Year Study design	1. Patient Selection*				2. Index Test*				3. Reference Standard*				4. Flow and Timing*			
	a. sample	b. exclusion	c. low risk of bias	d. applicable	a. reference results not known	b. threshold	c. low risk of bias	d. applicable	a. likely to classify	b. index results not known	c. low risk of bias	d. applicable	a. interval	b. same standard	c. all included in analysis	d. low risk of bias
Buhling, 2004 PCS	Yes	Yes	Yes	U	Yes	U	U	U	U	No	U	U	Yes	Yes	No	No
Cetin, 1996 PCS	U	Yes	U	No	Yes	Yes	Yes	Yes	Yes	U	U	Yes	Yes	Yes	Yes	Yes
Chastang, 2003 PCS	No	No	No	U	No	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes
Chevalier, 2011 PCS	Yes	Yes	Yes	U	Yes	Yes	Yes	Yes	Yes	U	U	Yes	Yes	No	Yes	No
De los Monteros, 1999 PCS	Yes	Yes	Yes	U	Yes	Yes	Yes	Yes	Yes	U	U	Yes	Yes	Yes	No	No
Deerochanawong, 1996 PCS	U	Yes	U	No	Yes	Yes	Yes	Yes	U	U	U	Yes	Yes	No	Yes	No
Eslamian, 2008 PCS	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	U	Yes	Yes	Yes	Yes
Gandevani, 2011 PCS	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	U	Yes	No	Yes	Yes
Hill, 2005 PCS	Yes	No	No	No	U	Yes	U	Yes	Yes	U	U	Yes	Yes	Yes	No	No
Jakobi, 2003 PCS	U	No	No	No	U	No	U	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Jensen, 2004 PCS	Yes	Yes	Yes	U	Yes	Yes	Yes	U	Yes	Yes	Yes	Yes	U	Yes	U	U
Kashi, 2007 PCS	No	No	No	No	Yes	Yes	Yes	U	Yes	No	No	Yes	U	Yes	No	No
Kauffman, 2006 PCS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lamar, 1999 PCS	Yes	U	U	Yes	Yes	Yes	Yes	Yes	Yes	U	U	Yes	Yes	Yes	No	No

Author, Year Study design	1. Patient Selection*				2. Index Test*				3. Reference Standard*				4. Flow and Timing*			
	a. sample	b. exclusion	c. low risk of bias	d. applicable	a. reference results not known	b. threshold	c. low risk of bias	d. applicable	a. likely to classify	b. index results not known	c. low risk of bias	d. applicable	a. interval	b. same standard	c. all included in analysis	d. low risk of bias
Maegawa, 2003 PCS	Yes	Yes	Yes	U	Yes	Yes	Yes	Yes	Yes	U	U	U	Yes	No	Yes	No
Mello, 2006 PCS	Yes	Yes	Yes	U	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	No	Yes	No
Moses, 2011 PCS	Yes	Yes	Yes	U	Yes	Yes	Yes	Yes	U	Yes	U	U	Yes	Yes	Yes	Yes
Ostlund, 2003 PCS	Yes	Yes	Yes	U	Yes	Yes	Yes	Yes	Yes	No	No	Yes	U	Yes	U	No
Perea-Carrasco, 2002 PCS	Yes	Yes	Yes	U	U	Yes	U	Yes	Yes	U	U	Yes	U	Yes	Yes	U
Perucchini, 1999 PCS	Yes	Yes	Yes	U	Yes	Yes	Yes	U	Yes	U	U	Yes	Yes	Yes	U	Yes
Poyhonen-Alho, 2004 PCS	Yes	Yes	Yes	U	Yes	Yes	Yes	Yes	U	No	Yes	U	Yes	No	U	U
Rajput, 2012 PCS	Yes	Yes	Yes	No	No	No	No	Yes	Yes	U	U	Yes	Yes	Yes	Yes	Yes
Reichelt, 1998 PCS	Yes	U	U	No	U	No	U	U	Yes	U	U	Yes	U	Yes	U	U
Rey, 2004 PCS	Yes	U	U	Yes	Yes	Yes	Yes	Yes	Yes	U	U	Yes	U	Yes	Yes	U
Rust, 1998 PCS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	U	U	Yes	Yes	No	Yes	Yes
Sacks, 2003 PCS	Yes	U	U	Yes	Yes	Yes	Yes	Yes	Yes	U	U	Yes	Yes	Yes	Yes	Yes
Siribaddana, 2003 PCS	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Soheilykhah, 2011 PCS	Yes	Yes	Yes	No	No	Yes	No	Yes	U	No	No	Yes	U	Yes	U	U

Author, Year Study design	1. Patient Selection*				2. Index Test*				3. Reference Standard*				4. Flow and Timing*			
	a. sample	b. exclusion	c. low risk of bias	d. applicable	a. reference results not known	b. threshold	c. low risk of bias	d. applicable	a. likely to classify	b. index results not known	c. low risk of bias	d. applicable	a. interval	b. same standard	c. all included in analysis	d. low risk of bias
Soonthornpun, 2003 PCS	No	U	No	No	U	Yes	U	Yes	Yes	U	U	Yes	Yes	Yes	Yes	Yes
Tan, 2007 PCS	Yes	Yes	Yes	No	U	Yes	U	Yes	Yes	U	U	Yes	Yes	No	Yes	No
Trihospital 1998 PCS (Naylor)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Uncu, 1995 PCS	U	U	U	No	Yes	Yes	U	Yes	Yes	U	U	Yes	Yes	Yes	Yes	Yes
van Leeuwen 2007 PCS	Yes	Yes	Yes	U	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	U	Yes	U
Weerakiet, 2006 PCS	Yes	Yes	Yes	No	U	No	No	U	Yes	U	U	Yes	Yes	Yes	Yes	Yes
Wijeyaratne, 2006 PCS	Yes	U	U	No	Yes	U	U	U	Yes	Yes	Yes	Yes	Yes	U	Yes	U
Yachi, 2011 PCS	U	Yes	U	U	Yes	Yes	Yes	U	Yes	U	U	U	No	No	Yes	Yes
Yogev, 2004 PCS	U	Yes	U	Yes	Yes	Yes	Yes	Yes	Yes	No	U	Yes	Yes	U	Yes	U

*QUADAS domain descriptions: 1.a. Random or consecutive sample; 1.b. Did the study avoid inappropriate exclusions?; 1.c. Was the study a low risk of bias?; 1.d.Is the study applicable?; 2.a. Reference standard results not known; 2.b. Pre specified threshold; 2.c. Was the study a low risk of bias?; 2.d. Is the study applicable?; 3.a. Likely to classify target patients; 3.b. Index test results not known; 3.c. Was the study a low risk of bias?; 3.d. Is the study applicable?; 4.a. Interval between tests; 4.b. Same standard for all patients; 4.c. All patients included in analysis; 4.d. Was the study a low risk of bias?
U = unclear

Table C2. Methodological quality of randomized controlled trials (RCTs) using the Cochrane Collaboration's tool for assessing risk of bias for Key Questions 2 to 5

Author Year	Sequence generation	Allocation concealment	Blinding		Incomplete outcome data*	Selective outcome reporting	Other	Overall Risk of Bias* (quality rating)†
			Participants*	Outcome assessment*				
Bevier, 1999	Unclear	Unclear	Unclear	Low	Unclear	Low	Low	Unclear (fair)
Bonomo, 2005	Low	Unclear	Unclear	Low	Low	Low	Low	Unclear (fair)
Crowther, 2005	Low	Low	Low	Low	Low	Low	Low	Low (good)
Garner, 1997	Unclear	Unclear	Unclear	High	High	Low	Low	High (poor)
Landon, 2009	Unclear	Unclear	Low	Low	Low	Low	Low	Unclear (fair)

* Domains for which assessments are made by outcome were assessed for objective outcomes
† Quality rating based on EPC Methods Guide (good, fair, poor)

Table C3. Methodological quality of prospective cohort studies (PCS) and retrospective cohort studies (RCS) using Newcastle-Ottawa Quality Assessment Scale, by Key Question and design

Author, Year	Representative-ness of exposed cohort	Selection of non-exposed cohort	Ascertainment of exposure	Outcome of interest not present at start of study	Comparability of cohorts *(Study controls)*		Assessment of outcome	Follow up long enough for outcomes to occur	Adequacy of cohort follow up	Total stars (quality rating)[†]
					*known factors**	*additional factor*				
KQ2 - PCS										
Solomon, 1996	Selected group of users	Same community as exposed cohort	Structured interview	Yes	No	No	Record linkage	Yes	Subjects lost unlikely to introduce bias	6 (fair)
KQ2 - RCS										
Chanprapaph, 2004	Truly representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Complete follow up	7 (good)
KQ3 – PCS										
Ardawi, 2000	Truly representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Subjects lost unlikely to introduce bias	7 (good)
Lao, 2001	Somewhat representative	Same community as exposed cohort	Secure record	Yes	No	No	No description	Yes	Subjects lost unlikely to introduce bias	6 (fair)
Lapolla, 2007	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	No description	Yes	Complete follow up	8 (good)
Metzger/ HAPO, 2008	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Independent blind assessment	Yes	Subjects lost unlikely to introduce bias	9 (good)
Retnakaran, 2008	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	No description	Yes	Subjects lost unlikely to introduce bias	8 (good)
Sacks, 1995	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	No description	Yes	Subjects lost unlikely to introduce bias	8 (good)
Sermer, 1995 RCT	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Subjects lost unlikely to introduce bias	9 (good)

Author, Year	Representative-ness of exposed cohort	Selection of non-exposed cohort	Ascertainment of exposure	Outcome of interest not present at start of study	Comparability of cohorts *(Study controls)*		Assessment of outcome	Follow up long enough for outcomes to occur	Adequacy of cohort follow up	Total stars (quality rating)[†]
					*known factors**	*additional factor*				
Shirazian, 2008	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	No description	Yes	Subjects lost unlikely to introduce bias	8 (good)
KQ3 - RCS										
Aberg, 2001	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
Adams, 1998	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
Berggren, 2011	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Subjects lost unlikely to introduce bias	9 (good)
Berkus, 1995	Somewhat representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Complete follow up	7 (good)
Biri, 2009	Truly representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Complete follow up	7 (good)
Black, 2010	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
Bo, 2004	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
Cheng, 2009	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
Chico, 2005	Somewhat representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Complete follow up	7 (good)
Chou, 2010	Somewhat representative	Same community as	Secure record	Yes	No	No	Record linkage	Yes	Complete follow up	7 (good)

Author, Year	Representative-ness of exposed cohort	Selection of non-exposed cohort	Ascertainment of exposure	Outcome of interest not present at start of study	Comparability of cohorts *(Study controls)*		Assessment of outcome	Follow up long enough for outcomes to occur	Adequacy of cohort follow up	Total stars (quality rating)[†]
					*known factors**	*additional factor*				
		exposed cohort								
Cok, 2011	Somewhat representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Complete follow up	7 (good)
Corrado, 2009	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Subjects lost unlikely to introduce bias	9 (good)
Hillier, 2007	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Independent blind assessment	Yes	Subjects lost unlikely to introduce bias	9 (good)
Jensen, 2002	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Subjects lost unlikely to introduce bias	9 (good)
Kim, 2002	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Subjects lost unlikely to introduce bias	8 (good)
Kwik, 2007	Truly representative	Same community as exposed cohort cohort	Secure record	Yes	No	No	Record linkage	Yes	Subjects lost unlikely to introduce bias	7 (good)
Langer, 2005	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
Lao, 2003	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	No	Record linkage	Yes	Subjects lost unlikely to introduce bias	8 (good)
Lapolla, 2011	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
Morikawa, 2010	Somewhat representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Complete follow up	7 (good)

Author, Year	Representative-ness of exposed cohort	Selection of non-exposed cohort	Ascertainment of exposure	Outcome of interest not present at start of study	Comparability of cohorts *(Study controls)*		Assessment of outcome	Follow up long enough for outcomes to occur	Adequacy of cohort follow up	Total stars (quality rating)[†]
					*known factors**	*additional factor*				
Nord, 1995	Somewhat representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Subjects lost unlikely to introduce bias	7 (good)
Pennison, 2001	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
Ricart, 2005	Truly representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Complete follow up	7 (good)
Rust, 1996	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Subjects lost unlikely to introduce bias	9 (good)
Schwartz, 1999	Truly representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Follow up rate <75% and no description of those lost	6 (fair)
Stamilio, 2004	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
Tan, 2008	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
Vambergue, 2000	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
Yang, 2002	Somewhat representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)
KQ4/5 - PCS										
Malcolm, 2006	Somewhat representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Follow up rate <75% and no description of those lost	6 (fair)

Author, Year	Representative-ness of exposed cohort	Selection of non-exposed cohort	Ascertainment of exposure	Outcome of interest not present at start of study	Comparability of cohorts *(Study controls)*		Assessment of outcome	Follow up long enough for outcomes to occur	Adequacy of cohort follow up	Total stars (quality rating) †
					*known factors**	*additional factor*				
KQ4/5 – RCS										
Adams, 1998	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Subjects lost unlikely to introduce bias	9 (good)
Bonomo, 1997	Selected group of users	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	8 (good)
Fassett, 2007	Somewhat representative	Same community as exposed cohort	Secure record	Yes	No	No	Record linkage	Yes	Complete follow up	7 (good)
Langer, 2005	Truly representative	Same community as exposed cohort	Secure record	Yes	Yes	Yes	Record linkage	Yes	Complete follow up	9 (good)

* Controls for known factors: age, race, BMI, history of GDM, family history of DM

† Quality rating based on EPC Methods Guide (good, fair, poor): total scores of 7-9 were considered good, 4-6 fair, and 0-3 poor.

BMI = body mass index; GDM = gestational diabetes mellitus; DM = diabetes mellitus; PCS = prospective cohort study; RCS = retrospective cohort study

Appendix D. Evidence Tables

Table D1. Characteristics of studies examining properties of current screening and diagnostic tests for GDM, Key Question 1

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Agarwal, 2000 June 1998 to Apr 2000 United Arab Emirates	1644 (+hx = 1276, +GCT 398) 29.8±5.87 (+hx) 30.2±5.62 (+GCT) NR	**Inclusion:** attending antenatal clinic; referred for OGTT because of clinical history or +OGCT **Exclusion:** pre-screened by other methods	Selective, 2-step CC, 513/1644 (31.2%) +hx, 396/1276 (31.0%) +GCT, 117/368 (31.8%) Risk factor: anytime during pregnancy OGCT: 24-28 wks	FPG, various thresholds (taken same time as OGTT) 28.1 wks (+hx) 28.7 wks (+GCT)	CC, 1991 100g, 3 h 1-2 wks after OGCT	**Purpose:** Investigate the value of FPG as an alternative screen to OGCT **Recommendations:** In a high-risk population FPG offers a simple and practical screening test
Agarwal, 2001 Dec 1997 to May 1998 United Arab Emirates	430 (HbA1c) 426 (cFruc) NR NR	**Inclusion:** attending antenatal clinic; referred for OGTT **Exclusion:** NR	Selective, 2-step Risk factor OGCT ADA, 116/430 (27.0%)	HbA1c ≥5.0% cFruc ≥210 µmol/L	ADA, 1997/CC, 1991 100g, 3 h 1-2 wks after OGCT/ 24-28 wks risk factor screen	**Purpose:** Investigate practical alternative screening amongst high-risk population which can be easily performed on a single blood sample **Recommendations:** Screening high-risk pregnancies with a combination of cFRUC and HbA1c could avoid OGTT in 37.9% women.

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Agarwal, 2005(a) May 2003 to Jul 2003 United Arab Emirates	442 **G1**: 26.2 ± 5.3 **G2**: 28.5 ± 5.9 NR	**Inclusion:** Attended routine antenatal clinics, 24-28 wks gestation, complete OGTT record **Exclusion:** Delivery in other hospital, failure to undergo OGTT, hepatic, renal or evident DM, diet treatment, previous GDM, any endocrine disorder	Universal, 1-step ADA, 85 (19%) WHO, 49 (11%) No screen	HbA1c (cutoff value ≥7.5%; collected at time of OGTT)	ADA, 2004 WHO, 1999 75 g, 2 h 24-28wks	**Purpose:** Is HbA1c is an effective screen for GDM **Recommendations:** HbA1c is a poor test to screen for GDM
Agarwal, 2005(b) Jun 2003 to Jan 2004 United Arab Emirates	1,685 26.6 ± 5.7 (non-GDM) 29.3 ± 6.4 (GDM) 27.7 ± 8.5 (non-GDM) 28.9 ± 5.6 (GDM)	**Inclusion:** Attended routine antenatal clinics at hospital, 24-28 wks gestation, complete OGTT record **Exclusion:** Delivery in other hospital, failure to undergo OGTT, hepatic, renal or evident DM, diet treatment, previous GDM, any endocrine	Universal, 1-step WHO, 333 (19.8%) No screen	FPG, <4.7 and >5.6 mmol/L	WHO, 1999 75 g, 2 h 24 to 28 wks	**Purpose:** evaluate the value of FPG in screening a high-risk population for GDM **Recommendations:** FPG has the potential to avoid nearly 1/3 of OGTTs at the expense of missing 1/5 of pregnant women with milder GDM

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Agarwal, 2006 May 2004 to Sep 2005 United Arab Emirates	4,602 28.4 ± 6 NR	**Inclusion:** Routine antenatal clinic attendance at study hospital **Exclusion:** NR	Universal, 2-step ADA, 675 (14.7%) WHO, 979 (21.3%) ADIPS, 1158 (25.2%) EASD, 556 (12.1%) 24-28 wks	FPG (various cutoff values)	ADA, 2004 WHO, 1999 ADIPS, 1999 EASD, 1998 75 g, 2 h 24-28 wks	**Purpose:** Effect of diagnostic criteria on the usefulness of FPG as a screen for GDM **Recommendations:** Initial testing by FPG can decrease the number of OGTTs needed to diagnose GDM
Agarwal, 2008 Nov 2006 to Jun 2007 United Arab Emirates	1,662 28.8 ± 5.9 NR	**Inclusion:** Routine antenatal clinic attendance **Exclusion:** NR	Universal, 1-step ADA, 186 (11.2%) No screen	FBG (hand-held glucometer; cutoff value ≥4.9 mmol/L)	ADA, 2004 75 g, 2 h 24-28 wks	**Purpose:** Test the practical value of measuring FBG vs. FPG **Recommendations:** FBG is a simple, practical, cost-effective and patient-friendly approach to screen for GDM
Agarwal, 2011 Oct 2008 to May 2009 United Arab Emirates	849 29.4 ± 6.0 NR	**Inclusion:** Routine antenatal care **Exclusion:** Twin pregnancy, pregestational DM, Hx of GDM	Universal, 1-step ADA, 113 (13.3%) IADPSG, 279 (32.9%) WHO, 156 (20.3%) ADIPS, 172 (20.3% EASD, 90 (10.6%) 24-28 wks	Serum fructosamine (cutoff value ≥237 µmol/L)	ADA, 2004 IADPSG, 2010 WHO, 1999 ADIPS, 1999 EASD, 1998 75 g, 2 h 24-28 wks	**Purpose:** Evaluate the value of serum fructosamine to screen for GDM **Recommendations:** Serum fructosamine is a poor test to screen for GDM

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Ardawi, 2000 Jun 1996 to Jun 1998 Saudi Arabia	818 **G1**:29.2 ± 4.6 **G2**:30.7 ± 4.8 **G3**:32.1 ± 5.1 NR	**Inclusion:** Attended antenatal care clinics at 2 hospitals **Exclusion:** NR	Universal, 2-step NDDG, 102 (12.5%)	50 g OGCT (≥7.2 mmol/L)	NDDG, 1979 100 g, 3 h 24-28 wks	**Purpose:** Evaluate applicability of the 50 g OGCT as a screening test for GDM in relation to pregnancy outcomes **Recommendations:** 50 g OGT at 24-28 weeks with a cutoff value of 7.8 mmol/L is a reliable screening test for GDM
Ayach, 2006 Jul 1997 to Dec 1999 Brazil	341 Age ≥25, n = 54 (15.8%) BMI ≥27, n = 49 (14.4%)	**Inclusion:** All pregnant women, no Hx of DM, sought care in study hospital during 1st half of pregnancy **Exclusion:** Failure to perform or finish screening/diagnostic test, withdrawal of consent or premature termination of pregnancy, miscarriage, pseudocyesis, premature birth, fetal death, intolerance to oral glucose test	Universal, 2-step ADA, 13 (3.8%) 24-28 wks	FPG and risk factors (age ≥ 25, BMI before pregnancy ≥ 27 kg/m^2, family or personal history of diabetes, and membership of an ethnic group with high prevalence of GDM)	ADA, 2002 100 g, 3 h 24-28 wks	**Purpose:** Compare FPG + risk factors vs.50 g GTT **Recommendations:** FPG + risk factors are more appropriate for screening compared with 50 g OGCT

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Balaji, 2011 NR India	1,463 23.6 ± 3.3 21.5 ± 4.1	**Inclusion:** Visiting antenatal clinic for the first time in second or third trimester **Exclusion:** Hx of GDM or DM	Universal, 1-step WHO, 196 (13.4%) No screen	FPG (IADPSG ≥5.1 mmol/L) 24-28 wks	WHO, 1999 75 g, 2 h 24-28 wks	**Purpose:** Ascertain the ability of FPG to diagnose glucose intolerance during pregnancy in Asian Indians **Recommendations:** FPG is not suitable for diagnosis of GDM in this population
Balaji, 2012 NR India	819 23.8 ± 3.48 21.2 ± 4.87	**Inclusion:** Pregnant women at 24-28 wks, attending community health center **Exclusion:** NR	Universal, 1-step WHO, 86 (10.5%) No screen	CBG (point-of-care testing with glucometer; 75 g glucose load, 2 h sample, cutoff value of ≥7.8 mmol/L)	WHO, 1999 75 g, 2 h 24-28 wks	**Purpose:** Compare point-of-care measured CBG with a glucometer and lab-estimated VPG **Recommendations:** CBG value at a 2 h plasma glucose ≥7.8 mmol/L may be recommended for the diagnosis of GDM
Berkus, 1995 NR U.S.	80 **G1**: 28.1 ± 5 **G2**: 25.7 ± 5 NR	**Inclusion:** Non-hypertensive women, recruited from obstetric clinic in Texas, non diabetic **Exclusion:** NR	NR, 2-step NDDG, 21/40 (26%) WHO, 20/40 (50%)	50 g OGTT 75 g OGTT, WHO	NDDG, 1979 100g, 3 h **G1**:28.6 ±4 **G2**:30.6 ±4	**Purpose:** Determine whether glucose abnormality, as shown by GTT periodicity, is not affected by different glucose loads **Recommendations:** GTT periodicity identifies patients with GDM regardless of GTT load

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Bobrowski, 1996 Jul 1992 to Jan 1994 U.S.	422 NR	**Inclusion:** + OGCT screen **Exclusion:** no follow-up OGTT	NR (included women with abnormal OGCT) 24-28 wks NDDG, 124(29%) CC, 161 (38%)	50 g OGTT, ≥135 mg/dL	NDDG, 1979 CC,1982 100 g, 3 h 1-2 wks after GCT	**Purpose:** examine the utility of various 50 g screen cutoff values in establishing the diagnosis of gestational diabetes **Recommendations:** 50-g glucose screen result ≥220 mg/dL can obviate the need for a 3-h OGTT
Brustman, 1995 NR U.S.	32 28 ± 5 NR	**Inclusion:** Women 26-26 wks gestation, abnormal glucose screen ≥130 mg/dl after 24 wks gestation **Exclusion:** NR	NR (included women with abnormal OGCT) NDDG, 16 (50%) IWC, 6 (19%)	IWC, 3rd (75 g, 3-h OGTT)	NDDG, 1979 100 g, 3 h 26-36 wks	**Purpose:** Compare results of a 75 g, 3h OGTT with a 100g OGTT **Recommendations:** 75 g OGTT using the NDDG criteria, recognizes carbohydrate intolerance in pregnancy
Buhling, 2004 Jun 1997 to Jan 2000 Germany	912 28.5 ± 5 23.6 ± 4.4	**Inclusion:** Received prenatal care at clinic, no previous GDM testing **Exclusion:** NR	Universal, 2-step ADA, 37 (4.1%)	50g OGCT (≥140 mg/dL)	ADA, 2001 75 g, 2 h 33.8 ±3 wks	**Purpose:** Evaluate the sensitivity of the glucose-sticks for screening for GDM **Recommendations:** Urine glucose dip stick analysis is not useful to detect GDM

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Cetin, 1996 Oct 1994 to Jan 1996 Turkey	274 G1:27 (19-37) G2: 28 (18-37) G3: 29 (19-41) G1: 24.8 (17.3-40.1) G2: 24.5 (17-40) G3: 25 (19.3-39.8)	**Inclusion:** Women > 24 yrs, 24-28 wks gestation, examined by obstetrician before 20 wks, singleton pregnancy **Exclusion:** Hx of preexisting diabetes, preeclampsia, regular ingestion of any drug, delivery ≤28 wks , premature rupture of membranes	Universal, 2-step NDDG, 17 (6.2%)	50g, 1 h OGCT (≥140 mg/dL)	NDDG, NR 100 g, 3 h 26-28 wks	**Purpose:** Examine different cutoff values with regard to the time of patient's last meal **Recommendations:** Different cutoff values lead to improved efficiency of the OGCT and decreased frequency of OGTT
Chastang, 2003 Jun 1997 to Jun 1998 France	354 31.4 ± 4.6 22.5 ± 4.1	**Inclusion:** Presented at least 1 RF for GDM: >35 years, BMI > 25, family Hx of diabetes, personal Hx of GDM, Hx of macrosomia/ LGA, Hx or preeclampsia, presence of obstetrical event(s) in current pregnancy, excessive weight gain during in current pregnancy **Exclusion:** NR	Selective, 2-step CC, 69 (20%) 24 to 28 wks	≥25 g carbohydrate breakfast FPG	CNGOF, 1998 (based on CC criteria) 100 g, 3 h 24-28 wks	**Purpose:** Validate a diagnostic test for GDM which predicts the risk of macrosomia **Recommendations:** Standard 50 g carbohydrate breakfast is more sensitive than the 50 g GCT to screen women at risk of macrosomia

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Chevalier, 2011 Jan 2002 to Dec 2006 France	11,545 32.8 ± 5.5 (GDM) 30.7 ± 5.3 (no GDM) 28.6 ± 5.7 (GDM) 27.8 ± 4.9 (no GDM)	**Inclusion:** screened between 24-28 at hospital **Exclusion:** NR	Universal, 2-step CC (≥130 mg/dL), 344 (4.3%) CC (≥140 mg/dL), 300 (3.9%) 24-28 wks	OGCT, ≥130 mg/dL and ≥140 mg/dL	CNGOF, 1996 (based on CC, 1982) 100 g, 3 h	**Purpose:** Explore GDM screening according to the 1996 French guidelines **Recommendations:** Two-step screening strategy for GDM was neither relevant nor efficient
De Los Monteros, 1999 Jul 1996 to Dec 1996 Mexico	445 >25 (n=359) <25 (n=86) NR	**Inclusion:** 24-28 wks gestation, attending medical centre for routine care **Exclusion:** Previous Hx of DM, consent withdrawal during either glucose tolerance test, inability to recall last menstrual period, Hx of regular drug ingestion during pregnancy	Universal, 2-step NDDG, 43 (9.7%) CC, 52 (11.7%) Sacks, 62 (13.9%) 24-28 wks	Postprandial 50 g OGCT	NDDG, 1979 CC, 1982 Sacks, 1989 100 g, 3 h 1 wk after OGCT	**Purpose:** Study sensitivity and specificity of the 50 g, 1 h GCT performed 1 to 2 h after a non-standardized home breakfast **Recommendations:** Sensitivity after breakfast was similar, based on the NDDG and CC criteria for GDM
Deerochana-wong, 1996 NR Thailand	709 26.9 ± 5.6 22.4 ± 3.8	**Inclusion:** Attending antenatal clinic, no prepregnancy DM **Exclusion:** NR	Universal, 2-step NDDG, 10 (1.4%) WHO, 111 (15.7%) 24-28 wks	50 g OGCT WHO, 1980 (75 g, 2 h OGTT)	NDDG, 1979 100 g, 3 h Within 7 days	**Purpose:** Compare criteria of the NDDG and WHO for pregnancy outcomes **Recommendations:** WHO criteria resulted in poorer pregnancy outcomes but fewer perinatal complications were missed than with the NDDG criteria

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Eslamian, 2008 NR Iran	138 27.5 ± 4.6 24.9 ± 3.1	**Inclusion:** Patients receiving prenatal care **Exclusion:** Pre-gestational DM, current GDM	Universal, 2-step CC, 12 (8.6%) 24-28 wks	Standard breakfast containing 50 g simple sugar	CC, NR 100 g, 3 h	**Purpose:** Compare a standard breakfast with a 50 g glucola-based OGCT **Recommendations:** Standard breakfast can be used as an alternative method for assessing carbohydrate intolerance
Gandevani, 2011 2007 to 2008 Iran	1,804 32.5 ± NR 23.3 ± 2.4	**Inclusion:** Prenatal clinic attendance at study center, referred for 50 g GCT between 24-28 wks **Exclusion:** Glucose intolerance before pregnancy, Hx of GDM	Universal, 2-step CC, 130 (7.2%)	50 g OGCT (various cutoff values)	CC, 1982 100 g, 3 h 24-28 wks	**Purpose:** Investigate cutoff value of GCT in an Iranian population **Recommendations:** Best cutoff value is 135 mg/dL to identify GDM
Hill, 2005 Jun 1997 to Aug 1998 India	830 24 (16-40) 23.1 (20.7-25.7)	**Inclusion:** Planned to deliver at hospital, singleton pregnancy, <32 wks GA determined by LMP or a first trimester ultrasound scan **Exclusion:** Prepregnancy DM	Selective, 2-step CC, 49 (6%) NR	Risk Factors (one or more of the following: BMI ≥25 kg/m^2; family Hx of DM in a first or second degree relative; poor obstetric Hx; previous baby weighing ≥3800 g; PIH; polyhydramnios	CC, 1982 100 g, 3 h 28-32 wks	**Purpose:** Determine the incidence of GDM in one urban maternity unit in South India and examine its effect on the offspring **Recommendations:** Effect of maternal glucose concentrations on neonatal anthropometry is continuous and extends into those diagnosed as normal

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Jakobi, 2003 1998 to 1999 Israel	180 NR NR	**Inclusion:** Positive 50 g OGCT (≥7.8 mmol/L), referred to high-risk pregnancy clinic **Exclusion:** NR	Abnormal OGCT IWC, 25 (13.9%) NR	BG/ portable glucose meter	IWC, 3rd (similar to NDDG) 100 g, 3 h 28-32 wks	**Purpose:** Evaluate perinatal effects of replacing current methods for 100 g OGTT with portable glucose meters **Recommendations:** No difference between the 2 methods
Jensen, 2003 1999 to 2000 Denmark	5,235 NR NR	**Inclusion:** Risk group: women presenting ≥1 RF. Non-risk group: contacted by study midwife at first appointment **Exclusion:** Preexisting DM, <18 yrs, delivery or migration before 30 wks, first booking later than 30 wks	Universal, 2-step WHO, 124 (2%) NR	Risk factors (glucosuria, GDM in a previous pregnancy, prepregnancy BMI ≥27 kg/m^2, family history of DM, and previous delivery of macrosomic infant)	WHO, 1998 75 g, 2 h 28-32 wks	**Purpose:** Evaluate a screening model for GDM using clinical risk indicators **Recommendations:** Using risk factor assessment reduces the need for screening and diagnostic testing in 66% pregnant women

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Kashi, 2007 NR Iran	200 27.8 ± 95.2 29.6 ± 4.5	**Inclusion:** Referred to prenatal clinic, ≥1 risk factors: >25 years old, Hx of recurrent abortion, previous GDM, preeclampsia, macrosomia, still birth, DM in first degree family or pregestational BMI >25 kg/m^2 **Exclusion:** Pregestational overt DM	Selective, 2-step ADA, 20 (10%) 24-28 wks	FPG (≥91.5 mg/dL)	ADA, 2006 100 g, 3 h 1-2 wks after +OGCT	**Purpose:** Determine a cutoff point of FPG for screening for GDM **Recommendations:** FPG level of 91.5 mmol/dL showed highest sensitivity and specificity
Kauffman, 2006 NR U.S.	123 NR NR	**Inclusion:** Women attending obstetrical clinic, 24-28 wks gestation with consent to undergo 100 g, 3h OGTT in lieu of 50 g screen **Exclusion:** Hx DM or GDM, untreated endocrine disorders, medications with impact on circulating glucose or insulin levels	No screen, OGTT in lieu of OGCT NDDG, 16 (13.0%) CC, 25 (20.3%)	homeostatic insulin sensitivity indices (HOMA-1, HOMA-2, QUICKI) FPG ≥92 mg/dL FPI ≥93 µmol/L	NDDG, 1979 CC, 1982 100 g, 3 h 24-28 wks	**Purpose:** investigate homeostatic indices of insulin sensitivity to screen for GDM **Recommendations:** FPG and the homeostatic insulin sensitivity indices are sensitive alternatives to OGCT

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Lamar, 1999 NR U.S.	136 26 ± 5.3 NR NDDG, 5 (3.7%)	**Inclusion:** Women in general obstetric population at institution ≥18 yrs and between 24-28 wks, no Hx of overt DM **Exclusion:** NR	NR, 2-step NDDG, 5 (3.7%) 24-28 wks	50 g OGCT (traditional and alternative sugar source - 28 jelly beans consisting of 50 g of simple sugar)	ACOG, 1994 (Values same as NDDG) 100 g, 3 h Within 7-10 days of OGCT	**Purpose:** Determine if a standardized dose of jelly beans is an alternative sugar source to the 50 g glucose beverage to screen for GDM **Recommendations:** Jelly beans provide a "dose" of simple carbohydrate similar to that of the 50 g glucose beverage but with suboptimal sensitivity
Maegawa, 2003 Apr 1999 to Sep 2001 Japan	749 28.9±4.1 (normal) 30.7±2.5 (early) 34.1±3.3 (late) 21.0±2.9 (normal) 24.8±6.2 (early) 22.6±2.3 (late)	**Inclusion:** Women in 1st trimester; attending hospital **Exclusion:** Hx of DM	Universal, 2-step JSOG, 22 (2.9%)	GCT*, 130 mg/dL and140 mg/dL FPG* 85 mg/dL HbA1c *4,8% and 5.8% *Taken in both 1st and 2nd trimester	JSOG, 2002 (Values same as ADA, 75 g) 75 g, 2 h 2-4 wks after 2nd trimester screen	**Purpose** Characteristics of various screening procedures for GDM in Japan during the first trimester and between 24 and 28 wks of pregnancy **Recommendations:** Of 22 with GDM, 14 were diagnosed in the first trimester and 8 in the second trimester.
Mello, 2006 Jan 1997 to Dec 1999 Italy	227 (16-20 wks) 976 (26-30 wks) NR NR	**Inclusion:** nonobese; nondiabetic; singleton pregnancy **Exclusion:** NR	Universal, 1-step Early: CC 41/227 (18.1%) ADA 15/227 (6.75) Late: CC 60/484 (12.4%) ADA, 26/484 (4.4%)	75 g, 2 h (ADA) OGTT 16-20 wks 26-30 wks	CC, 1982 100 g, 3 h 1 wk after 75 g	**Purpose:** Investigate the comparability of the 75 g and the 100 g tests in the diagnosis of GDM **Recommendations:** There was only weak diagnostic agreement between 75-g and 100-g glucose loads

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Moses, 2011 Jan 2010 to Jun 2010 Australia	1 ,275 NR NR	**Inclusion:** NR **Exclusion:** NR	Univresal, 1-step ADIPS, 123 (9.6%) IADPSG, 166 (13.0%)	IADPSG, 2010	ADIPS, 1991 75 g, 2 h NR	**Purpose:** Compare the prevalence of GDM using IADPSG criteria vs. ADIP criteria **Recommendations:** IADPSG criteria Increased the prevalence of GDM from 9.6% to 13.0%
Ostlund, 2003 Jul 1994 to Jun 1996 Sweden	4,918 NR NR	**Inclusion:** Nondiabetic women visiting maternal health care clinics in Sweden **Exclusion:** Pre-pregnancy DM	Universal, 2-step WHO, 61 (1.7%) NR	Anamnestic risk factors (Heredity, non-Nordic origin, prior macrosomia, prior GDM, multipara, prior macrosomia, and prior GDM)	WHO, 1980 75 g, 2 h 28-32 wks	**Purpose:** Determine prevalence of GDM and the value of traditional anamnestic risk factors for predicting the outcome of the OGTT **Recommendations:** Traditional risk factors as an indicator to perform an OGTT gives a low sensitivity to detect GDM

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Perea-Carrasco, 2002 NR Spain	578 NR NR	**Inclusion:** Attended routine antenatal clinic, OGCT and OGTT between 24-28 wks **Exclusion:** Multiple pregnancies	Universal, 2-step IWC, 46 (7%) 24-28 wks	Index test (I) = (fructosamine/ total protein) - (glucose/100) I = ≥27.2	IWC, 3rd (same as NDDG thresholds) 100 g, 3 h 24-28 wks	**Purpose:** Devise an index test to improve screening sensitivity and specificity, offering better screening capability and greater ease of diagnosis **Recommendations:** Proposed index offers an efficient screening test for GDM, and with more stringent cutoff points may be applicable as a single-step diagnostic procedure
Perucchini, 1999 1995 to 1997 Switzerland	520 28.4 ± 0.2 23.8 ± 0.2	**Inclusion:** Singleton pregnancy, attended hospital, delivery >28 wks **Exclusion:** Pre-existing DM, not examined before 24 wks	Universal, 2-step IWC, 53 (10.2%)	FPG (≥4.8 mmol/L, 86 mg/dL)	IWC, 4th (similar to CC/ADA 2000/10) 100 g, 3 h 24-28wks	**Purpose:** Evaluate FPG vs. the 1 h 50 g OGCT **Recommendations:** More women are referred for the OTT using FPG vs. those using the OGCT
Poyhonen-Alho, 2004 Jan 1996 to Aug 1998 Finland	532 NR NR	**Inclusion:** Caucasian, attendance at primary health care units **Exclusion:** Pre-pregnancy DM	Universal, 2-step Author defined, 123 (23%)	Risk factor based screening (BMI >27; age >40; previous child >4500 g; previous GDM; glucosuria; or macrosomia in current pregnancy)	Author defined 75 g, 2 h Fasting ≥4.8, 1h ≥10.0, 2h ≥8.7 mmol/L 26-28 wks	**Purpose:** Compare whether universal screening by OGCT will identify more women with GDM vs. risk factor based screening **Recommendations:** 50 g OGCT identified a higher number women with GDM

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Rajput, 2011 NR India	607 16–20 18%; 21–25 58%; 26–30 20%; >30 4% <18.5 38%; 18.5–24.9 54%; ≥25 8%	**Inclusion:** all pregnant women 24-28 wks GA; **Exclusion:** know Dx DM, anemia, chronic renal, pancreatic or other severe illness	Universal, 1-step ADA, 43 (7.1%) IADPSG, 144 (23.7%)	HbA1c , >5.45% and >5.25% (diagnostic)	ADA, 2010 IADPSG, 2010 75 g, 2 h 24-28 wks	**Purpose:** Evaluate the utility of HbA1c in combination with OGTT for diagnosis of GDM **Recommendations:** HbA1c in combination with OGTT can obviate the need of OGTT in almost two-thirds of women with GDM
Reichelt, 1998 May 1991 to Aug 1995 Brazil	4,977 27.9 ± 5.5 26.1 ± 4.1	**Inclusion:** Women ≥20 yrs, 21-28 wks gestation **Exclusion:** Pre-pregnancy DM	Universal, 1-step WHO, 379 (7.6%)	FPG (≥87 mg/dL)	WHO, 1994 75 g, 2 h 24-28 wks	**Purpose:** Evaluate FPG as a screening test for GDM **Recommendations:** FPG is a useful screening test for GDM
Rey, 2004 9 mo period Canada	188 30.2 ± 5.2 NR	**Inclusion:** all women between 24 and 28 wks; normal first-trimester glucose testing; screened according to CDA screening program **Exclusion:** NR	Normal 1st trimester screen CDA, 21 (11.2%) 25.7 ± 1.2 wks	GCT, 7.8 mmol/L FPG, 4.5 mmol/L FCG, 4.6 mmol/L	CDA, 1998 75 g, 2 h 27.2 ± 1.4 wks	**Purpose:** compare the performance in screening of the 1 h, 50 g GCT, FPG and FCG **Recommendations:** There is not enough benefit to be gained by using the FPG instead of the GCT as the screening test for GDM

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Rust, 1998 Jul 1994 to Jun 1995 U.S.	448 23.7 ± 6.1 26.8 ± 7.6	**Inclusion:** Women at medical centre obstetric clinics, >20 wks gestation **Exclusion:** NR	Universal, 2-step ADA, 16 (3.6%) ≥20 wks	Postprandial 50 g GCT (1 and 2 hrs post meal glucose load)	ADA, 100 g, 3 h 20 wks	**Purpose:** Compare 2 h postprandial glucose measurements with the 1 h, 50 g glucola screen as a predictor of GDM **Recommendations:** 1 h glucola test is a reliable screening test for GDM whereas the 2 h postprandial test is not
Sacks, 2003 Feb 1998 to Jul 1999 U.S.	4,507 NR NR	**Inclusion:** Prenatal visit at medical center, no known diabetic Hx, able to return for lab work and glucose testing **Exclusion:** Transferred care to other institution, began prenatal care or screened elsewhere, spontaneous abortion after enrollment	Universal, 2-step ADA, 302 (6.7%) ≥23 wk	FPG (≥83 mg/dL)	ADA, 2001 75 g, 2 h NR	**Purpose:** Determine whether the FPG test administered at the first prenatal visit is an efficient screen for GDM **Recommendations:** FPG has poor specificity (high false-positive rate) making it an inefficient screening test

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Siribaddana, 2003 NR Sri Lanka	721 NR NR	**Inclusion:** Attended antenatal clinic hospital **Exclusion:** Known DM	Universal, 2-step WHO, 40 (5.5%) 24-28 wks	50g OGCT Traditional risk factors (Age, family Hx, parity, Hx of poor pregnancy outcomes)	WHO, 1985 75 g, 2 h 1 wk after OGCT	**Purpose:** Determine the prevalence of GDM in a Sri Lankan population using WHO criteria, and establish the predictive value of a 50g OGCT vs. the OGTT **Recommendations:** Traditional risk factors did not predict GDM; screening for GDM should be performed in all women with a GCT
Soheilykhah, 2010 2007 to 2010 Iran	1,502 27.3 ± 6.1 25.7 ± 6.9	**Inclusion:** Attended prenatal clinics **Exclusion:** Hx hyperglycemia, on medication known to affect glucose metabolism	Universal, 2-step ADA, 216 (13.1%) 24-28 wks	Time intervals of 100 g OGTT	ADA, 2009 100 g, 3 h 1-2 wks after +OGCT	**Purpose:** To find an appropriate and simple way to perform screening tests for GDM **Recommendations:** A positive GCT result (≥130 mg/dL) with subsequent 2 h 100g OGTT (≥150 mg/dL) will diagnose GDM

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Soonthornpun, 2003 NR Thailand	42 33.6 ± 5.4 NR	**Inclusion:** 50 g OGCT values ≥140 mg/dL at screening between 14-36 wks **Exclusion:** NR	NR (included women with abnormal OGCT) CC, 9 (21.4%) ADA, 3 (7.1%)	ADA, 2000 (75 g, 2 h GTT)	CC, 1982 100 g, 3 h 28.2 ± 4.2	**Purpose:** Test the validity of a 75 g, 2 h OGTT using the ADA criteria and reference values for the 100 g, 3 h OGTT **Recommendations:** The prevalence of GDM was lower using the 75 g OGTT using the criteria and reference values of the 100 g OGTT
Tan, 2007 Jan 2006 to Jul 2006 Malaysia	521 29.6 ± 4.8 26.7 ± 4.6	**Inclusion:** antenatal booking; ≥1 risk factors **Exclusion:** NR	Universal, 2-step Selective, 2-step WHO, 180 (34.5%) 28.8 ± 6.4 wks	Clinical risk factors, 1 or more: ≥35 years, Hx macrosomia ≥4 kg; Hx intrauterine death; weight ≥70 kg, BMI ≥30, Hx of GDM, family Hx DM, or glycosuria	WHO, 1999 75 g, 1 h	**Purpose:** Evaluate the role of risk factors in conjunction with GCT to determine an appropriate threshold for 1 h GCT **Recommendations:** 2-step screening threshold for a positive GCT should be ≥ 7.6 mmol/L. After a GCT result, clinical risk factors are no longer useful in selecting women.

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Tri-Hospital (2 papers) Sermer, 1998 Naylor, 1997 Sept 1989 to Mar 1992 Canada	3,836 NR NR	**Inclusion:** >24 yrs at time of delivery, no Hx of DM examined by physician before 24 wks gestation, delivery >28 wks **Exclusion:** NR	Universal, 2-step NDDG, 145 (3.8%) 26-28 wks	50 g, 2 h OGCT (time of last meal prior to glucose load)	NDDG, 1979 100 g, 3 h 26-28 wks	**Purpose:** Established more efficient screening strategies for detection of GDM **Recommendations:** Increasing maternal carbohydrate intolerance is associated with a graded increase in adverse maternal and fetal outcomes
Uncu, 1995 NR Turkey	42 27.5 ± 4.3 NR	**Inclusion:** Attending outpatient clinic, OGCT between 24-28 wks **Exclusion:** Pregnancies beyond wk 28 previously diagnosed as DM	Universal, 2-step CC, 14 (33%) 24-28 wks	Serum fructosamine (≥2.85 mmol/L) HbA1c (≥7.2%)	CC, 1988 100 g, 3 h NR	**Purpose:** Evaluated the sensitivity and specificity of 50 g OGCT, serum fructosamine and HbA1c levels as screening tests for GDM **Recommendations:** HbA1c and fructosamine levels are reliable methods to 50 g OGCT
van Leeuwen, 2007 NR Netherlands	1,301 30.8 ± 4.9 24.2 ± 4.6	**Inclusion:** NR **Exclusion:** Known preexisting diabetes; no prenatal care before 24 wks of gestation	Universal, 2-step WHO, 48 (3.7%) 24-28 wks	Random 50 g glucose test	WHO, NR 75 g, 2 h NR	**Purpose:** Compare the accuracy measures of the random glucose test and the 50 g GCT as screening tests for GDM **Recommendations:** The 50 g glucose challenge test is more useful than the random glucose test

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD (kg/m^2)*	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Weerakiet, 2006 Jul 2004 to Mar 2005 Thailand	359 31.8 ± 6.1 23.2 ± 4.3	**Inclusion:** Singleton pregnancy, presenting ≥1 risk factor for GDM: age >30, obesity, family Hx of DM, prior GDM, glucosuria, signs of hyperglycemia, Hx of poor obstetric outcome **Exclusion:** Hypertension, known DM, known chronic disease requiring Tx, positive result for syphilis, hepatitis B (HBSAg), HIV	Selective, 2-step ADA, 66 (16.7%) Risk factor screen recommended by ACOG 21-27 weeks (OGCT)	Adiponectin levels (10 µmg/mL) 50g OGCT (≥140 mg/dL)	ADA, 2000 100 g, 3 h 24-28 wks	**Purpose:** Evaluate adiponectin as a predictive factor for GDM and appropriate as a screening test for GDM **Recommendations:** Adiponectin was not as strong a predictor as GCT
Wijeyaratne, 2006 Apr 2003 to Jul 2003 Sri Lanka	853 NR NR	**Inclusion:** Registered for antenatal care **Exclusion:** Established glucose intolerance	Selective, 2-step WHO, 144 (16.3%) 24-28 wks	FBG (≥4.1 mmol/L) FPG (≥4.7 mmol/L) Risk factors proposed by ADA, NR	WHO, 1999 75 g, 2 h 24-28 wks	**Purpose:** Evaluate tests used for screening and confirmation of GDM in Sri Lanka **Recommendations:** Urine and FBG are unsuitable for screening

Author, year Dates of study Country	Women Analyzed, *n* Maternal Age, *mean ± SD/median ± IQR (yr)* BMI, *mean ± SD* (kg/m^2)	Inclusion/Exclusion Criteria	Screening Practice^ *Prevalence of GDM Criteria, n (%)* Time of Screening	Index†, (Comment)	Reference†*, Date Load, Interval Time of GDM Confirmation	Study Purpose Conclusion(s)
Yachi, 2011 Sep 2008 to Jan 2010 Japan	509 33.4 ± 3.7 20 ± 2.5	**Inclusion:** Visited clinic; ≥13wks gestation **Exclusion:** FPG levels ≥2.5 mmol/L; missing or incomplete data	Universal, 2-step JSOG, 8 (2.0%) 24-29 wks	FPG (≥3.66 mmol/L at 10 wks) FPI (≥36.69 mmol/L at 10 wks)	JSOG, 1999 75 g, 2 h 26-29 wks	**Purpose:** Determine early screening tests and risk factors predictive of glucose intolerance in later pregnancy **Recommendations:** FPG is not an acceptable screening test for glucose intolerance
Yogev, 2004 1995 to 1999 U.S.	2,541 26.1±6.3(>130) 29.2±7.0 (>180) 26.4±3.7 (>130) 27.6±3.1 (>180)	**Inclusion:** Singleton pregnancies; screened at 24-28 wks **Exclusion:** No Hx of GDM and pre-gestational DM	Universal, 2-step CC, 469 (6.8%) NDDG, NR (7.3%) 24-28 wks	50 g OGCT (130, 135, 140 mg/dL)	CC, 1982 NDDG, 1979 100 g, 3 h +OGCT only, 1-2 wks OGCT	**Purpose:** Describe the predictive value for GDM using different OGCT thresholds in Mexican-American women **Recommendations:** A threshold of ≥130 mg/dL is recommended

Notes: ^ Screening practice described in study;†Index and reference data used in this review. *Complete diagnostic criteria can be found in Table 1. ADA = American Diabetes Association; ADIPS = Australian Diabetes in Pregnancy Society; BMI = body mass index; CBG = capillary blood glucose; CHT = chronic hypertension; d = day; dL = deciliter; DM = diabetes mellitus; Dx = diagnosis/diagnostic; EASD = European Association for the Study of Diabetes FPG = fasting plasma glucose; GA = gestational age; GCI = gestational carbohydrate intolerance; GCT = glucose tolerance test; GDM = gestational diabetes mellitus; GHT = gestational hypertension; HbA1c = glycated hemoglobin; HBSAg = hepatitis B virus surface antigen; HOMA = homeostatic model assessment; h = hour; mg = milligrams; IADPSG = International Association of Diabetes and Pregnancy Study Groups; IWC = International Workshop Conference; JSOG = Japan Society of Obstetrics and Gynecology; CNGOF = National College of French Obstetricians and Gynaecologists; NDDG = National Diabetes Data Group; NR = not reported; OGTT = oral glucose tolerance test; PCS = prospective cohort study; PIH = pregnancy-induced hypertension; PROM = premature rupture of the membrane; QUICKI = Quantitative insulin sensitivity check index ; RCS = retrospective cohort study; RF = risk factors; SD = standard deviation; Tx = treatment; WHO = World Health Organization; wk(s) = week(s); yr(s) = year(s)

Table D2. Characteristics of studies comparing outcomes for women who were and were not screened for GDM, Key Question 2

Author, year Study Design, Duration of Followup Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD (kg/m^s)*	Inclusion/ Exclusion Criteria	Gestational Age at Screening Screening Test	Outcomes Reported Numbers screened vs. not screened, *n (%)*
Chanprapaph, 2004 RCS, Until birth Thailand	1,000 Screened: 31.5 ± 5.5 Not screened: 24.0 ± 3.8 Screened: 22.5 ± 3.8 Not screened: 20.9 ± 2.9	**Inclusion:** Pregnant women attending a single antenatal care center; attendance from Oct 2001 to Dec 2002. **Exclusion:** NR	First booking; 24 & 28 wks; 30 & 32 wks **Step 1:** Risk factors + 50 g OGCT; positive ≥ 140 mg/dL after 1 hour **Step 2:** 100 g OGTT: 1) fasting glucose value 105 mg/dL 2) 1 hr 190 mg/dL 3) 2 hr 165 mg/dL 4) 3 hr 145 mg/dL -test considered positive if any 2 of non-fasting values greater than normal	**Obstetric complications:** PROM: 30 (7) vs. 46 (8) PIH: 21 (5) vs. 7 (1) GHT: 4 (1) vs. 4 (1) CHT: 4 (1) vs. 2 (0.3) PPH: 3 (1) vs. 1 (0.2) Chorioamnionitis: 0 (0) vs. 1 (0.2) Polyhydramnios: 1 (0.2) vs. 0 (0) Total obstetric complications: 65 (16) vs. 63 (11) **Pregnancy outcomes:** Preterm delivery: 42 (10) vs. 50 (8) Birthweight: >90th percentile: 50 (12) vs. 55 (9) <10th percentile: 42(10) vs. 58 (10) Fetal anomalies: 3 (2) vs. 1 (1) Cesarean section: 81 (20) vs. 71 (12)
Solomon, 1996 RCS, Until birth US	93 Screened: 30.5 Not screened: 31.1 Screened: 23.0 Not screened: 23.6	**Inclusion:** Female nurses; 25 to 42 yrs residing in 1 of 14 US states **Exclusion:** NR	**Gestational Age:** NR **Step 1:** 1 h 50 g OGCT	**Maternal morbidity:** NR **Fetal morbidity:** Macrosomia (7% each group)

* BMI = body mass index; CHT = chronic hypertension; dl = deciliter; DM = diabetes mellitus; GA = gestational age; OGCT = oral glucose tolerance test; GDM = gestational diabetes mellitus; GHT = gestational hypertension; mg = milligrams; NR = not reported; OGTT = oral glucose tolerance test; PCS = prospective cohort study; PIH = pregnancy – induced hypertension; PROM = premature rupture of the membrane; RCS = retrospective cohort study; SD = standard deviation; wk = weeks; yr = years

Table D3. Characteristics of studies examining outcomes of mothers and offspring in the absence of treatment, Key Question 3

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Aberg, 2001 RCS (4) Sweden Jan 1995 - Dec 1997	4,657 **G1:** Sub-GDM Group (no Tx) **G2:** Control (no Tx)	NR NR	**Inclusion:** Singleton pregnancy, within Lund University hospital register, results matched **Exclusion:** NR	2 h, 75 g OGTT WHO, NR	Emergency cesarean delivery, elective cesarean delivery, perinatal mortality rate **Other:** Gestational duration, birth weight, umbilical artery pH, APGAR score
Adams,1998 RCS (1) US Jan 1986 - Sep 1996	389 **G1:** GDM Diet (Tx) **G2:** GDM Insulin (Tx) **G3:** Unrecognized GDM (no Tx) **G4:** Control (no Tx)	**G1:** 31.4 ± 4.9 **G2:** 31.5 ± 4.6 **G3:** 30.2 ± 4.7 **G4:** 30.2 ± 4.5 **G1:** 26.1 ± 6.1 **G2:** 30.3 ± 7.2 **G3:** 26.6 ± 7.5 **G4:** 26.3 ± 7.0	**Inclusion:** Positive OGCT; meets NDDG criteria (2 plasma glucose values on OGTT) for GDM **Exclusion:** Multiple gestation; fetal congenital anomalies; delivery before 34 wks; delivery elsewhere; diet or insulin therapy initiated < 4 wks before delivery	1 h, 50 g OGCT 3 h, 100 g OGTT NDDG, 1979	Cesarean delivery, maternal weight gain, maternal birth trauma (rectal injury), macrosomia (BW >4000 gm, >4500 gm), shoulder dystocia, clavicular fracture, brachial plexus injury (cranial nerve palsy, brachial plexus, permanent & healed), hypoglycemia, hyperbilirubinemia (within neonatal complications composite), mortality (stillbirth) **Other:** Birthweight, LGA, vacuum & forceps delivery
Ardawi, 2000 PCS (2) Saudi Arabia Jun 1996 – Jun 1998	818 **G1:** Negative Screenees (no Tx) **G2:** Positive Screenees (no Tx) **G3:** GDM by NDDG (Tx)	**G1:** 29.2 ± 4.6 **G2:** 30.7 ± 4.8 **G3:** 32.1 ± 5.1 **G1:** 64.3 ± 4.1 **G2:** 68.6 ± 4.1 **G3:** 75.2 ± 4.5	**Inclusion:** NR **Exclusion:** Hepatic renal disease, DM prior to pregnancy, previous diet therapy, previous GDM, known endocrine disorders	1 h, 50 g OGCT 3 h, 100 g OGTT NDDG 1979	Cesarean delivery, macrosomia, hypoglycemia, hyperbilirubinemia, mortality(stillbirth) **Other:** Fetal length, <25g, head circumference, wk at delivery

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Berggren, 2001 RCS (1) US Apr 1996 – May 2010	3,759 **G1**: CC GDM (no Tx) **G2:** NDDG GDM (Tx) **G3:** Control (no Tx)	NR NR	**Inclusion:** Delivery at UNC women's hospital **Exclusion:** No results available on 1 hr 50 g OGCT, delivery <24 wks, pregestational DM, GDM diagnosed by 50 g OGCT only	1 h, 50 g OGCT 3 h, 100 g OGTT NDDG 1979 CC 1982	Preeclampsia, Maternal Hypertension, Cesarean delivery, maternal birth trauma (3rd or 4th degree laceration), macrosomia, shoulder dystocia **Other:** GA at delivery, mode of delivery other than c-section, HELLP (hemolysis, elevated liver enzymes, low platelet count), birthweight, NICU admission, NICU stay >48 hrs
Berkus, 1995 RCS (NR) US 1987 – 1988	833 **G1:** GDM by CC (no Tx) **G2:** GDM by Sacks (no Tx) **G3:** GDM by Langer (no Tx) **G4:** Normal (no Tx)	**G1:** 29.0 ± 5.0 **G2:** 30.0 ± 7.0 **G3:** 29.0 ± 6.0 **G4:** 26.0 ± 6.0 NR	**Inclusion:** Nonhypertensive gravidas; singleton pregnancy; underwent 3-hour GTT; attended clinics in San Antonio area **Exclusion:** Women with 2+ abnormal OGTT values by NDDG criteria	No OGCT 3 h, 100 g OGTT Coustan & Lewis, 1978 NDDG, 1979 Langer,1987 Sacks,1989	Macrosomia **Other:** Birthweight
Biri, 2009 RCS (1) Turkey Jan 2004 - Dec 2006	2,029 **G1:** Normal 50 g GLT (no Tx) **G2:** Abnormal 50 g/ Normal 100 g (no Tx) **G3:** 1 Abnormal 100 g (no Tx) **G4:** GDM - 100 g GLT (Tx) **G5:** GDM – 50 g GLT (no Tx)	**G1:** 29.6 ± 4.6 **G2:** 30.9 ± 4.9 **G3:** 32.1 ± 4.6 **G4:** 33.3 ± 4.8 **G5:** 32.6 ± 5.0 NR	**Inclusion:** Singleton pregnancies, screened at study centre **Exclusion:** Prepregnancy DM, multiple gestations	1 h, 50 g OGCT 3 h, 100 g OGTT ACOG, 2001 NDDG, 1979	Preeclampsia, cesarean delivery, macrosomia, hypoglycemia, hyperbilirubinemia **Other:** Birthweight, LGA/SGA, APGAR, respiratory complications, polyhydramnios, prematurity

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Black, 2010 RCS (1) US Oct 2005 – Mar 2010	8,711 All no Tx **G1:** No GDM **G2:** IGT **G3:** IFG **G4:** IGT-2 **G5:** IFG-IGT	**G1:** 28.6 ± 5.9 **G2:** 32.1 ± 5.4 **G3:** 30.4 ± 5.6 **G4:** 32.3 ± 5.2 **G5:** 32.0 ± 5.1 **G1:** 26.9 ± 5.8 **G2:** 28.1 ± 5.6 **G3:** 30.8 ± 7.1 **G4:** 27.5 ± 4.7 **G5:** 31.8 ± 7.0	**Inclusion:** Singleton birth >20 wks gestation, received 2 hr 75 g OGTT with no prior 50 g OGCT, available pre-pregnancy and delivery anthropometric data **Exclusion:** Any form of treatment	2 h, 75 g OGTT IADPSG, 2010	Cesarean delivery, maternal weight gain, gestational hypertension, shoulder dystocia/birth injury, hyperbilirubinemia **Other:** Birthweight, LGA, ponderal index, preterm delivery
Bo, 2004 RCS (1) Italy Apr 1999 - Feb 2001	700 **G1:** OGCT negative (normal) (no Tx) **G2:** OGCT positive OGTT negative (no Tx) **G3:** OGTT1 abnormal value (Tx) **G4:** GDM positive (Tx)	**G1:** 30.8 ± 4.2 **G2:** 31.8 ± 4.3 **G3:** 32.9 ± 4.7 **G4:** 32.6 ± 4.9 NR	**Inclusion:** Caucasian; attending clinic **Exclusion:** Known DM, any disease affecting glucose metabolism	1 h, 50 g OGCT 3 h, 100 g OGTT CC, 1982	Cesarean delivery, macrosomia, hyperbilirubinemia (icterus), mortality (death) **Other:** "Metabolic Syndrome in Pregnancy", premature births, birthweight, LGA/SGA, APGAR score, respiratory distress, malformations, neonatal diseases
Cheng, 2009 RCS (1) US Jan 1988 - Dec 2001	1,469 **G1:** No GDM (no Tx) **G2:** GDM by CC only (no Tx) **G3:** GDM NDDG only (Tx)	NR NR	**Inclusion:** All pregnancies screened and delivered at University of California **Exclusion:** Multifetal pregnancies, vaginal breech deliveries, delivery <24 wks, congenital anomalies, pregestational DM	1 h, 50 g OGCT 3 h, 100 g OGTT NDDG,1979 CC, 1982	Preeclampsia, cesarean delivery (mode of delivery), maternal birth trauma (3rd or 4th degree laceration), macrosomia, shoulder dystocia, birth trauma composite variable incl. brachial plexus injury, facial nerve palsy, clavicular and skull fracture, head laceration **Other:** Preterm delivery <37wks, APGAR <7, neonatal acidemia, LGA

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Chico, 2005 RCS (1) Spain Jan 1999 - Dec 2001	6,248 **G1:** Standard criteria (Tx) **G2:** New criteria (Tx) **G3:** Subgroup- New IGT criteria (no Tx) **G4:** Normal tolerance (no Tx)	**G1:** 33.4 ± 4.0 **G2:** 33.3 ± 4.0 **G3:** 33.3 ± 4.0 **G4:** 32.8 ± 4.0 NR	**Inclusion:** All pregnancies handled in 2 yr period **Exclusion:** None	1 h, 50 g OGCT 3 h, 100 g OGTT NDDG, 1979 4th IWC/ADA, 2003 4th IWC/CC, 1998	Cesarean delivery, maternal weight gain, macrosomia (>4000 g), hypoglycemia, hyperbilirubinemia (jaundice), mortality (fetal deaths) **Other:** Week of delivery, instrumentation, birthweight, LGA/SGA, APGAR, malformations
Chou, 2010 RCS (1) Taiwan Jan 2001 - Sep 2008	10,990 **G1:** Normal (no Tx) **G2:** GDM by CC but not NDDG criteria (no Tx) **G3:** GDM by NDDG criteria (Tx)	**G1:** 32.8 ± NR **G2:** 33.4 ± NR **G3:** 34.4 ± NR NR	**Inclusion:** Singleton pregnancies delivered at Cathay General Hospital **Exclusion:** Multiple pregnancies, fetal anomalies diagnosed prenatally	1 h, 50 g OGCT 3 h, 100 g OGTT CC, 1982 NDDG, 1979	Maternal hypertension, cesarean delivery, maternal birth trauma (postpartum hemorrhage), macrosomia, shoulder dystocia, mortality (intrauterine fetal demise) **Other:** Preterm labour, APGAR scores
Cok, 2011 RCS(1) Turkey Jan 2003 - Jun 2009	185 **G1:** 0h OGTT (no Tx) **G2:** 1 h OGTT (no Tx) **G3:** 2 h OGTT (no Tx) **G4:** 3 h OGTT (no Tx)	**G1:** 32.5 ± 4.8 **G2:** 30.1 ± 4.5 **G3:** 30.0 ± 5.1 **G4:** 30.2 ± 4.3 **G1:** 33.7 ± 4.5 **G2:** 30.8 ± 3.8 **G3:** 29.8 ± 4.3 **G4:** 30.1 ± 3.2	**Inclusion:** Women presenting to Baskent Unviersity, one abnormal OGTT value **Exclusion:** Multiple gestations or prepregnancy DM, 2 abnormal OGTT values	1 h, 50 g OGCT 3 h, 100 g OGTT CC, 1982	Macrosomia **Other:** LGA, birthweight, birth week
Corrado, 2009 RCS (NR) Italy	776 **G1:** OAV (no Tx) **G2:** Control (no Tx)	**G1:** 31.2 ± 5.1 **G2:** 30.1 ± 4.9 **G1:** 25.0 ± 5.1 **G2:** 24.2 ± 4.4	**Inclusion:** Caucasian, one positive screening test and OGTT **Exclusion:** Multiple gestations, Tx for GDM	1 h, 50 g OGCT 3 h, 100 g OGTT CC, 1982	Preeclampsia/maternal hypertension (hypertensive disorders of pregnancy), cesarean delivery, macrosomia, hypoglycemia

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Jan 1996 - Dec 2005			(insulin/diet)		**Other:** GA, birthweight, APGAR
Hillier, 2007 RCS (2) US 1995-2000	9,439 **G1:** Normal (no Tx) **G2:** Positive OGCT normal OGTT (no Tx) **G3:** Positive OGCT and 1 Abnormal CC or NDDG (no Tx) **G4:** GDM-CC (no Tx), **G5:** GDM NDDG (Tx)	NR NR	**Inclusion:** Data on mother-child pairs 5-7 yrs PP **Exclusion:** Preexisting DM	1 h, 50 g OGCT 3 h, 100 g OGTT NDDG, NR (1979) CC criteria as presented in 4th IWC, 1998	Macrosomia (maternal glycemic level associated with macrosomia, childhood obesity) **Other:** Prevalence, risk of childhood obesity; association with maternal GDM screening results during pregnancy (hyperglycemia)
Jensen, 2002 RCS(4) Denmark Jan 1992 - Dec 1996	3,260 **G1:** Normal WHO (no Tx) **G2:** Normal DPSG but IGT WHO (no Tx) **G3:** Abnormal DPSG and IGT WHO (Tx) **G4:** GDM by both (Tx)	NR NR	**Inclusion:** First pregnancy in study period, tested with 75 g OGTT **Exclusion:** Pregestational GDM, multiple pregnancies, chronic disease	2 h, 75 g OGTT WHO, 1985 DPSG, 1991	Preeclampsia, maternal hypertension, cesarean delivery, maternal weight gain, macrosomia (>4000g), hypoglycemia, hyperbilirubinemia (jaundice) **Other:** LGA, respiratory distress, preterm delivery, glucosuria, GA
Kim, 2002 PCS(1) South Korea NR	699 **G1:** Normal (no elevated) **G2:** 1 Elevated (1 h elevated) **G3:** 2 Elevated (2 h elevated) **G4:** 3 Elevated (3 h elevated)	**G1:** 30.7 ± 3.9 **G2:** 29.5 ± 4.4 **G3:** 30.2 ± 3.3 **G4:** 32.3 ± 3.8 **G1:** 21.4 ± 2.9 **G2:** 21.0 ± 3.0 **G3:** 20.7 ± 2.6 **G4:** 21.8 ± 2.8	**Inclusion:** singleton pregnancy; antenatal care at Ajou University Hospital Department of Obstetrics and Gynecology **Exclusion:** missing data; confirmed GDM dx	1 h, 50 g OGCT 3 h, 100 g OGTT NDDG, NR	Preeclampsia, cesarean delivery, birthweight, LGA 90th percentile (macrosomia), hypoglycemia, perinatal death **Other:** Gestational age at birth (wks), APGAR, respiratory distress syndrome, poor perinatal outcome

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Kwik, 2007 RCS(1) Australia Feb 2000/Oct 2003 - May 2005	675 **G1:** Treated **G2:** Untreated **G3:** Comparison	**G1:** 34.5 ± 4.8 **G2:** 33.3 ± 4.7 **G3:** 32.8 ± 4.5 **G1:** 23.8 ± 4.4 **G2:** 22.9 ± 4.6 **G3:** 22.6 ± 3.7	**Inclusion:** Singleton pregnancy, 75 g GTT with a fasting value ≤ 5.5 mmol/L and 2-h blood sugar ≥7.8 mmol/L **Exclusion:** Confined ≤34 wks gestation	1 h, 50 g OGCT 2 h, 75 g OGTT ADA, 2000	Preeclampsia, cesarean delivery, macrosomia (BW > 4000 g), shoulder dystocia, clavicular fracture, brachial plexus injury (Erb's Palsy) **Other:** Mean birthweight, SCN admission, APGAR, premature delivery, GA at delivery
Landon, 2009 (primary) Landon, 2011 RCT(Multicenter, n = NR) US Oct 2002 - Nov 2007	1,841 **G1:** CC Mild GDM (no Tx) **G2:** CC False-positive, further divided by normal/ abnormal OGTT value (no Tx, no distinct data) **G3:** Normal control (no Tx)	**G1:** 28.9 ± 5.6 **G2:** 27.4 ± 5.5 **G3:** 25.1 ± 5.3 **G1:** 30.2 ± 5.1 **G2:** 30.1 ± 5.3 **G3:** 29.9 ± 5.8	**Inclusion:** Between 24 wks 0 ds and 30 wks 6 ds gestation, 135 and 200 mg/dL 1 hour after a 50 g glucose loading test **Exclusion:** Preexisting diabetes, abnormal results before 24 wks, prior GDM, Hx of stillbirth, multifetal gestation, asthma, CHT, corticosteriod use, known fetal anomaly, likely preterm delivery	1 h, 50 g OGCT 3 h, 100 g OGTT CC, 1982 4th IWC, 1998	Preeclampsia, maternal hypertension, cesarean delivery, maternal weight gain, macrosomia (BW >4000 g), shoulder dystocia, birth injury (trauma), hypoglycemia, hyperbilirubinemia, mortality (stillbirth/neonatal death) **Other:** GA at birth, elevated c-cord peptide, birthweight, LGA/SGA, Fat mass, Preterm delivery, NICU admission, IV glucose Tx, respiratory distress

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Langer, 2005 US(1) RCS Jan 1999 - Sept 1999	2,775 **G1:** GDM (no Tx, Dx after 37 wks) **G2:** GDM (Tx) **G3:** Nondiabetic control (no Tx)	**G1:** 27.6 ± 6.0 **G2:** 29.1 ± 6.0 **G3:** 25.0 ± 6.0 NR	**Inclusion:** Singleton pregnancies, FPG<140 mg/dL on OGTT; Case-control groups: GDM diagnosed >37 wks, treated GDM and non diabetic matched 2:1 for obesity, parity, ethnicity, GA at delivery (within 5 ds), yr of delivery **Exclusion:** Pregestational DM, substance abusers, multifetal gestation, fetal anomalies	1 h, 50 g OGCT 3 h, 100 g OGTT CC, 1982	Cesarean delivery, macrosomia, shoulder dystocia, hypoglycemia, hyperbilirubinemia, mortality (stillbirth) **Other:** Birthweight, LGA, ponderal Index >2.85, arterial cord pH <7.2, erythrocytosis, respiratory complication, induction of labour
Lao, 2001 PCS(1) China NR	487 **G1:** GDM by WHO (Tx) **G2:** Normal OGTT only (no Tx) **G3:** Control (no Tx)	**G1:** 32.1 ± 4.6 **G2:** 30.4 ± 5.3 **G3:** 27.7 ± 4.0 **G1:** 22.6 ± 3.2 **G2:** 22.0 ± 2.7 **G3:** 21.1 ± 2.7	**Inclusion:** Singleton pregnancies with visits to antenatal care between 28-30 wks **Exclusion:** Preexisting DM, CHT or other medical complication, thalassemia trait	2 h, 75 g OGTT WHO, 1980	Preeclampsia, cesarean delivery, maternal birth trauma (antepartum hemorrhage) **Other:** Preterm labor, prelabor rupture of the membranes, delivery mode, weeks gestation, birthweight, LGA/SGA, APGAR score 1 min., NICU admission
Lao, 2003 RCS(1) China 1996 – 1997	2,149 2 h OGTT (mmol/L): **G1:** <6.0 (no Tx) **G2:** 6.0 -6.9 (no Tx) **G3:** 7.0 -7.9 (no Tx)	**G1:** 28.6 ± 4.6 **G2:** 29.6 ± 4.6 **G3:** 30.8 ± 4.4 **G1:** 21.5 ± 2.6 **G2:** 21.7 ± 2.7 **G3:** 21.8 ± 2.8	**Inclusion:** Singleton pregnancy, antenatal OGTT, delivery at Queen Mary hospital, no insulin requirements **Exclusion:** Significant medical complications, taking no medication (ie. corticosteriods)	2 h, 75 g OGTT WHO, 1980	Cesarean delivery, macrosomia **Other:** Birthweight, LGA/SGA, preterm birth

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Lapolla, 2007 PCS(5) Italy NR	611 **G1:** Normal Control (no Tx) **G2:** False Positive (no Tx) **G3:** 1 Abnormal Glucose Value (OAV) (no Tx) **G4:** GDM (Tx)	**G1:** 30.9 ± 4.7 **G2:** 31.7 ± 4.9 **G3:** 32.5 ± 4.4 **G4:** 33.4 ± 4.4 **G1:** 22.4 ± 4.2 **G2:** 22.8 ± 3.9 **G3:** 23.7 ± 4.7 **G4:** 24.7 ± 4.8	**Inclusion:** No smoking; no CHT/specific conditions known to affect glucose metabolism **Exclusion:** Those with conditions known to affect glucose metabolism	1 h, 50 g OGCT 3 h, 100 g OGTT HbA1c* Criteria not defined, values same as Carpenter-Coustan	Cesarean delivery, macrosomia **Other:** LGA, ponderal index
Lapolla, 2011 RCS(1) Italy 1998 - 2008	1,927 **G1:** GDM formerly normal (no Tx) **G2:** Normal (no Tx)	**G1:** 32.4 ± 4.5 **G2:** 32.2 ± 4.5 **G1:** 23.7 ± 4.3 **G2:** 23.3 ± 4.2	**Inclusion:** Positive 50 g OGCT (1-h plasma glucose ≥ 7.8mmol/L), 3-h OGTT at 24–28 wks; negative result on OGCT or OGTT formed control group **Exclusion:** NR	1 h, 50 g OGCT 3 h, 100 g OGTT IADPSG, 2010 4th IWC, 1998	Maternal morbidity (eclampsia), maternal hypertension, cesarean delivery, macrosomia, shoulder dystocia (within fetal morbidity, incl. malformations, hypoglycemia,asphyxia,hyperbilirubinemia, etc.) **Other:** LGA/SGA, birthweight, ponderal index
Metzger/ HAPO, 2008 PCS(15) Various Jul 2000 - Apr 2006	23,316 (All no Tx) **G1:** 100 mg/dL + **G2:** 95-99 mg/dL **G3:** 90-94mg/dL **G4:** 85-89mg/dL **G5:** <85mg/dL; subdivided into **G6:** <75mg/dL **G7:** 75-79 mg/dL	**Tot:** 29.2 ± 5.8 **Tot:** 27.7 ± 5.1	**Inclusion:** Pregnant women **Exclusion:** <18 years, unknown LMP, no ultrasonographic estimation of GA between 6-24 wks, no OGTT within 32 wks, multiple pregnancies, assisted conception/IVF, glucose testing before recruitment, participation in another study or previous HAPO study, HIV, hepatitis B or C virus; no English language proficiency	2 h, 75 g OGTT HAPO Criteria; defined by groups	Preeclampsia, maternal hypertension, cesarean delivery, shoulder dystocia, hypoglycemia, hyperbilirubinemia **Other:** Cord blood serum C-peptide, Cord blood PG, CHT, intensive neonatal care, premature delivery, BW >90th percentile

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Morikawa, 2010 RCS(1) Japan Jan 2002- Dec 2006	228 **G1:** JSOG GDM (Tx) **G2:** JSOG - No GDM (no Tx) **G3:** IADPSG-Hyperglycemia (Tx) **G4:** IADPSG-New Patients (no Tx) **G5:** IADPSG No GDM (no Tx)	NR NR	**Inclusion:** Women with both OGCT and OGTT; singleton birth **Exclusion:** NR	1 h, 50 g OGCT 2 h, 75 g OGTT IADPSG, 2010 JSOG, 2008	Macrosomia **Other:** BW percentile
Nord, 1995 RCS(2) Sweden 1989 -1990	614 **G1:** 2-h OGTT 8.0-8.9mmol/L (no Tx) **G2:** Controls (no Tx)	**G1:** 30 ± 18-46 **G2:** 29 ± 16-45 **G1:** 22.3 ± 17.0 -43.3 **G2:** 21.3 ± 16.0-41.8	**Inclusion:** Intervention group: Indications to perform OGTT (Hx of DM in first degree relative; obesity (≥120 % or >9 kg); previous LFD-baby (>4.5 kg); IGT in previous pregnancy; accelerated fetal growth or polyhydraminosis; glucosuria; random B-glucose ≥7. mmol/L). Control group: No indication to perform OGTT **Exclusion:** NR	2 h, 75 g OGTT WHO, 1980	Preeclampsia, cesarean delivery, macrosomia (LFD - large for date), clavicular facture, brachial plexus injury, birth injury (traumatic delivery), hypoglycemia, hyperbilirubinemia, mortality **Other:** Premature delivery, respiratory distress syndrome, polycythemia requiring Tx, traumatic delivery

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Pennison, 2001 RCS (1) US 1995 - 1999	242 **G1:** Control (no Tx) **G2:** GDM NDDG (Tx) **G3:** GDM ADA (no Tx)	NR **G1:** 30.2 ± 1.1 **G2:** 31.7 ± 1.0 **G3:** 29.6 ± 1.1	**Inclusion:** Delivery at regional medical centre in Memphis; euglycemic or Dx GDM **Exclusion:** NR	1 h, 50 g OGCT 3 h, 100 g OGTT ADA, 1998 CC, 1982 NDDG/ACOG, 1994	Preeclampsia, cesarean delivery, macrosomia, shoulder dystocia, hypoglycemia
Retnakaran, 2008 PCS (Multicenter, n = NR) Canada 2003 - Sep 2007	396 **G1:** Normal OGCT, NGT (no Tx) **G2:** Abnormal OGCT, NGT (no Tx) **G3:** GIGT (no Tx) **G4:** GDM (Tx)	**G1:** 34.0 ± 4.4 **G2:** 33.8 ± 4.2 **G3:** 34.2 ± 4.2 **G4:** 34.5 ± 4.3 **G1:** 23.0 ± 21.5-26.1 **G2:** 23.5 ± 21.1-27.5 **G3:** 23.5 ± 21.8-27.7 **G4:** 25.0 ± 22.0-30.1	**Inclusion:** Attending outpatient obstetrics clinics; late second trimester; 50 g OGCT screen **Exclusion:** NR	1 h, 50 g OGCT 3 h, 100 g OGTT *3 mo. PP:* 2 h, 75 g OGTT NDDG 1979, CDA 2003	Maternal weight gain **Other:** 3 mo postpartum: maternal insulin sensitivity, beta-cell function, glycemia
Ricart, 2005 PCS (16) Spain 2002 - NR	9270 **G1:** NDDG GDM (Tx), **G2:** NDDG Negative (No Tx), **G3:** False-positive ADA (No Tx), **G4:** ADA GDM (No Tx)	**G1:** 31.9 ± 4.7 **G2:** 28.8 ± 5.3 **G3:** 30.5 ± 4.9 **G4:** 31.7 ± 4.6 **G1:** 25.9 ± 5.2 **G2:** 23.5 ± 3.9 **G3:** 24.5 ± 4.5 **G4:** 25.2 ± 4.7	**Inclusion:** Singleton pregnancy, no former Dx of GDM **Exclusion:** Women who did not undergo screening, unavailable results	1 h, 50g OGCT 3 h, 100g OGTT ADA, 2000 NDDG, 1979	Cesarean delivery, pregnancy induced hypertension, perinatal mortality, macrosomia **Other:** Preterm birth, LGA/SGA, APGAR score 1 & 5 mins, major malformations

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Rust, 1996 RCS(1) US NR - NR	664 **G1:** ≥ 2 of 4 values, abnormal by Sacks criteria **G2:** ≥ 2 of 4 values, abnormal by CC criteria **G3:** 1 abnormal by Sacks **G4:** 1 abnormal by CC **G5:** No abnormal by Sacks **G6:** No abnormal by CC	**G1:** 25.7 ± NR **G2:** 23.7 ± NR **G3:** 22.7 ± NR **G4:** 26.7 ± NR **G5:** 24.0 ± NR **G6:** 22.7 ± NR **G1:** 26.6 ± NR **G2:** 25.5 ± NR **G3:** 24.8 ± NR **G4:** 28.1 ± NR **G5:** 25.7 ± NR **G6:** 24.6 ± NR	**Inclusion:** Positive GDM screen result; underwent 3 h100 g OGT **Exclusion:** Delivery outside study hospital	1 h, 50 g OGCT 3 h, 100 g OGTT CC,1982 NDDG,1979 O'Sullivan and Mahan,1964 Sacks, 1975	Maternal hypertension,cesarean delivery, birth trauma (obstetric lacerations, hemorrhage) maternal weight gain, macrosomia, shoulder dystocia, birth trauma (dystocia disorders, birth trauma), hypoglycemia, hyperbilirubinemia, mortality (cumulative neonatal morbidity) **Other:** Intrauterine growth restriction, oligohydramnios, preterm labor, premature or prolonged rupture of the membranes, chorioamnionitis, malpresentation, labour induction, labour augmentation, fetal intolerance of labour, abdominal delivery, operative vaginal delivery
Sacks,1995 PCS(NR) US Mar 1992 - Mar 1993	3,505 **Groups:** Women were not grouped; actual glucose levels were used in regression analyses to assess the association with birthweight	**Tot:** 27.2 ± 5.8 **Tot:** 24.9 ± NR	**Inclusion:** Enrolled in prenatal care **Exclusion:** GDM in previous pregnancy, glucocorticoids, diet or insulin Tx, high fasting plasma glucose values, multiple gestations	2 h, 75 g OGTT No criteria defined, purpose of study to ID threshold values	Maternal weight gain, macrosomia

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Schwartz, 1999 RCS(4) US 1995 - 1996	8,711 **G1:** Normal results, prenatal screen (no Tx) **G2:** Abnormal (or no) prenatal screen and normal OGTT (no Tx) **G3:** NDDG GDM (Tx) **G4:** CC GDM (no Tx)	NR NR	**Inclusion:** No previous DM or GDM **Exclusion:** NR	1 h, 50 g OGCT 3 h, 100 g OGTT CC, 1982 NDDG, 1979	Cesarean delivery, macrosomia (BW >4000 g , >4500 g), mortality (stillbirth)
Sermer, 1995 (Primary) Naylor, 1996 RCT(3) Canada Sep 1989 - Mar 1992	3,780 **G1:** Negative screenees (no Tx), **G2:** False-positive Screenees (no Tx) **G3:** GDM- Borderline (no Tx) **G4:** GDM (Tx)	**G1:** 30.9 ± 4.1 **G2:** 31.9 ± 4.3 **G3:** 32.1 ± 4.4 **G4:** 32.7 ± 4.3 **G1:** 22.7 ± 3.8 **G2:** 23.1 ± 4.5 **G3:** 24.7 ± 5.8 **G4:** 24.2 ± 4.8	**Inclusion:** >24 yrs at delivery; no Hx of preexisting DM; examined by physician before 24 wks gestation **Exclusion:** Delivery <28 wks	1 h, 50 g OGCT 3 h,100 g OGTT NDDG, 1979 CC, 1982	Preeclampsia, cesarean delivery, macrosomia, hypoglycemia, hyperbilirubinemia (phototherapy) **Other:** Fetal trauma, congenital anomalies, respiratory distress syndrome, maternal/fetal length of stay
Shirazian, 2008 PCS(5) Iran NR - NR	612 **G1:** No GDM (no Tx) **G2:** GDM by ADA only (Tx) **G3:** GDM by WHO only (NR) **G4:** GDM by ADIPS only (NR)	NR Tot: 24.4 ± 4.6	**Inclusion:** No Hx of DM **Exclusion:** Pregestational DM, inablity to complete OGTT at 24-48 wks, twin pregnancies, no CHT, chronic renal failure, heart diseases, advanced pulmonary disease, current smokers, labor before 37th or after 40th gestational wk, planning to deliver at another hospital	2 h, 75 g OGTT ADA, 2008 WHO, 2008 ADIPS, 2008	Macrosomia

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Stamilio, 2004 RCS(1) US 1995 -1997	1,825 **G1:** False-positive OGCT (no Tx) **G2:** Negative OGCT (no Tx) **G3:** GDM (Tx)	**G1:** 28.5 ± NR **G2:** 25.5 ± NR **G1:** 28.5 ± NR **G2:** 25.5 ± NR	**Inclusion:** Delivery at University of Pennsylvania Medical Center, entry into triple marker screen perinatal database, complete followup **Exclusion:** Multiple gestations, anomalous fetuses	1 h, 50 g OGCT 3 h, 100 g OGTT NDDG modified by O'Sullivan cutoff, NR	Preeclampsia, maternal hypertension (chronic hypertension), long term hypertension (chronic hypertension), macrosomia, shoulder dystocia, mortality (antenatal death) **Other:** NICU admission, chorioamionitis, endometritis, birthweight (mean), high 28-week mean arterial pressure (maternal)
Tan, 2008 RCS(1) Malaysia Jan 2006 - July 2006	1,200 **G1:** Negative OGCT (no Tx) **G2:** False-Positive OGCT (no Tx)	**G1:** 28.9 ± 4.6 **G2:** 30.3 ± 4.7 **G1:** 26.5 ±4.4 **G2:** 27.0 ± 4.4	**Inclusion:** GCT screen at prenatal booking, GTT test only of GCT was positive, available delivery records **Exclusion:** Women missing GTT despite positive GCT, multiple gestations	1 h, 50 g OGCT 2 h, 75 g OGTT WHO, 1999	Cesarean delivery, maternal birth trauma (hemorrhage), macrosomia, SGA, fetal loss **Other:** Preterm birth, induction of labor, APGAR, cord blood ph
Vambergue, 2000 PCS(15) France Feb 1992 - Sep 1992	239 **G1:** Mild Gestational Hyperglycemia (MGH) (no Tx) **G2:** Control (no Tx)	**G1:** 28.8 ± 5.8 **G2:** 27.0 ± 5.2 **G1:** 24.8 ± 4.8 **G2:** 23.0 ± 3.9	**Inclusion:** Attendance at public maternity unit **Exclusion:** Twin pregnancies, pre-pregnancy high blood pressure, asthma, haemochromatosis, pre-pregnancy diabetes or GDM	1 h, 50 g OGCT 3 h, 100 g OGTT CC, 1982	Pregancy induced hypertension, cesarean delivery, shoulder dystocia, macrosomia, hypoglycemia, hyperbilirubinemia, mortality **Other:** LGA/SGA, respiratory distress, pathological deliveries, transfer to neonatal care unit, malformations, prematurity, APGAR score, adverse maternal and fetal outcome

Author, year Study Design (number of centers) Country Dates of study	Women Analyzed, *n* Groups	Maternal Age, *mean ± SD/ median ± IQR (yr)* BMI, *mean ± SD/ Median ± IQR (kg/m^s)*	Inclusion Criteria	Diagnostic Test Criteria	Outcomes Other Outcomes (Not defined by KQ)
Yang, 2002 PCS(16) China Dec 1998 - Dec 1999	404 **G1:** Impaired Glucose Tolerance (no Tx) **G2:** Normal (Normal Glucose Tolerance (no Tx)	**G1:** 28.0 ± 3.68 **G2:** 26.5 ± 2.95 **G1:** 22.6 ± 3.49 **G2:** 21.5 ± 2.57	**Inclusion:** NR **Exclusion:** <18 yrs, multiple pregnancies, maternal-fetal ABO incompatibility, maternal disease incl. prepregnancy diabeetes & those under long term medical treatment that may affect glucose metabolism, delivery outside Tianjin (rural or home delivery)	1 h, 50 g OGCT 2 h, 75 g OGTT WHO, 1998	Weight gain in pregnancy, cesarean delivery, birth trauma/dystocia, mild/moderate preeclampsia, birthweight > 90th percentile (macrosomia), birthweight > 95th percentile, hypoglycemia, perinatal death **Other:** PROM, breech presentation, preterm delivery, fetal male gender, low birth weight (< 2500 g), APGAR score < 7 @ 1 min, pneumonia

* ACOG = American Congress of Obstetricians and Gynecologists; ADA = American Diabetes Association; ADIPS = Australian Diabetes in Pregnancy Society; BMI = body mass index; CC = Carpenter-Coustan; CHT = chronic hypertension; d(s) = day(s); dL = deciliter; DM = diabetes mellitus; Dx = diagnosis/diagnostic; FPG = fasting plasma glucose; OGCT = oral glucose tolerance test; GDM = gestational diabetes mellitus; GLT = glucose load test; g = grams; HAPO = Hyperglycemia and Adverse Pregnancy Outcomes Study; h = hour; IADPSG = International Association of Diabetes and Pregnancy Study Groups; IFG = impaired fasting glucose; IGT = impaired glucose tolerance; IGT-2 = double impaired glucose tolerance; IQR = inter-quartile range; JSOG = Japan Society of Obstetrics and Gynecology; kg = kilogram; LGA = large for gestational age; m = meter; mg = milligrams; NDDG = National Diabetes Data Group; NR = not reported; OGTT = oral glucose tolerance test; PP= postpartum; PCS = prospective cohort study; PROM = premature rupture of the membranes; RCS = retrospective cohort study; SD = standard deviation; SGA = small for gestational age; Tx = treatment; wk(s) = week(s); WHO = World Health Organization; yr(s) = year(s)

Table D4. Characteristics of studies examining treatment outcomes of mothers and offspring, Key Questions 4 and 5

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
Adams,	389	**Inclusion:** Positive	**Screen:** 50 g GCT (24–30	**G1:** Diet with weekly	Weight gain,	NOS = 9

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
1998 RCS Jan 1986 to Sep 1996 US	**G1:** 31.5 ± 4.6 **G2:** 31.4 ± 4.9 **G3:** 30.2 ± 4.7 **G1:** 30.3 ± 7.2 **G2:** 26.1 ± 6.1 **G3:** 26.6 ± 7.5 NR **G1**: White: 73 **G2**: White: 277 **G3**: White: 15	OGCT; meets NDDG criteria (2 plasma glucose values on OGTT) for GDM **Exclusion:** Multiple gestation; fetal congenital anomalies; delivery before 34 wks; delivery elsewhere; diet or insulin therapy initiated < 4 wks before delivery	wks with 1-h cutoff by NDDG criteria, ≥ 140 mg/dL) **Diagnostic:**100 g OGTT at 24–30 wks (Fasting: 105 mg/dL;1 h 190 mg/dL; 2 h 165 mg/dL; 3 h 145 mg/dL)	blood glucose monitoring, daily BG self-monitoring and insulin required (n=76) **G2:** Diet with weekly blood glucose monitoring (n=297) **G3:** No treatment (n=16)	shoulder dystocia, hypoglycemia, stillbirth or neonatal death, birth trauma, birth weight, bone fracture/clavicular fracture, nerve palsy/brachial plexus injury, LGA, rectal injury, neonatal complications, Horner's syndrome, hemidiaphragm paralysis, unilateral eyelid ptosis from partial cranial nerve palsy	(good)

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
Bevier, 1999 RCT NR US	83 **G1:** 26.3 ± 6.0 **G2:** 27.4 ± 5.4 NR NR **G1**: White: 2 Black: 1 Hispanic: 45 **G2**: White: 2 Black: 0 Hispanic: 33	**Inclusion:** Positive OGCT screen and negative OGTT **Exclusion:** Hypertension; collagen disease; chronic renal disease; cardiac or pulmonary disease; Rh sensitization; Hx of preterm labor or SGA	**Screen:** 50 g GCT (24–30 wks with 1-h cutoff by NDDG criteria, ≥ 140 mg/dL) **Diagnostic:** 100 g OGTT (24–30 wks with fasting: 105 mg/dL; 1 h 190 mg/dL; 2 h 165 mg/dL; 3 h 145 mg/dL) HbA1c (28–32 wks)	**G1:** No diet, random glucose checks, and usual care (n=48) **G2:** Standard euglycemic diet, HBGM, random glucose checks HBGM recorded in a diary and reviewed weekly; 3 meals and 3 snacks: 40% carbohydrates, 20% protein, and 40% fat (n=35)	Preeclampsia, shoulder dystocia, birth weight, APGAR, abnormal fetal heart rate, SGA	RoB = Unclear (fair)

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
Bonomo, 1997 RCS 1989 to 1995 Italy	112 **G1:** 30.6 ± 3.4 **G2:** 30.7 ± 4.8 **G1:** 23.12 ± 4.4 **G2:** 25.0 ± 5.7 NR **G1**: NR **G2**: NR	**Inclusion:** Screened at diabetic centre; Dx of mild degree of glucose intolerance; OGCT >140 mg/dL and OAV on OGTT **Exclusion:** NR	**Screen:** 50 g GCT (14–16 wks for at risk and 24–28 wks for women without risk with 1 h cutoff) by CC and NDDG criteria **Diagnostic:** 100 g OGTT (14–16 wks for at risk and 24–28 wks for women without risk with Fasting, 1 h, 2 h, and 3 h intervals) by CC and NDDG criteria	**G1:** Elevated OGCT and Normal OGTT with no treatment from 1989 to 1993; from 1994 on patients given dietary advice; 25-30 kcal/kg per day diet; bi-weekly visits, BG monitoring (n=49) **G2:** 1 elevated OGTT with no treatment from 1989 to 1993; from 1994 on patients given dietary advice; 25-30 kcal/kg per day diet; bi-weekly visits, BG monitoring (n=63)	Caesarean delivery, birth weight, APGAR, LGA	NOS = 8 (good)

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
Bonomo, 2005 RCT 1997 to 2002 Italy	300 **G1:** 31.1 ± 4.7 **G2:** 30.7 ± 5.1 **G1:** 23.1 ± 4.4 **G2:** 23.0 ± 4.5 **G1:** fasting 4.68 ± 0.45 mmol/L **G2:** fasting 4.77 ± 0.52 mmol/L	**Inclusion:** Caucasian; OGCT >140 mg/dL and normal OGTT; singleton pregnancies **Exclusion:** Normal GCT; one abnormal OGTT value; GDM under CC criteria	**Screen:** 50 g GCT (24–28 wks with 1 h cutoff by Italian Society of Diabetology criteria, plasma glucose 1 h after challenge ≥ 7.8 mmol/L) **Diagnostic:**100 g OGTT (within 7 d of GCT) assessed by CC criteria GCT/OGTT repeated at 30–34 wks for complete diagnosis of Borderline Gestational Glucose Intolerance (BGGI)	**G1:** Diet and regular glucose monitoring; dietary counseling; 24–30 kcal/kg per day formal diet; caloric intake divided into 3 meals and 2–3 snacks; distributed as 50–55% carbohydrates, 25–30% protein, and 25% fat (n=150) **G2:** No special care, diet or treatment (n=150)	Caesarean delivery, weight gain, hypoglycemia, hyperbilirubinemia, admission to NICU, birth weight, weight, length, APGAR, LGA, ponderal index, SGA	RoB = Unclear (fair)

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
Chou, 2010 RCS (1) Jan 2001 to Sep 2008 Taiwan	10,990 **G1:** 34.4 ± NR **G2:** 33.4 ± NR **G1:** 23.11 ± NR **G2:** 23.45 ± NR NR NR	**Inclusion:** Singleton pregnancies delivered at Cathay General Hospital **Exclusion:** Multiple pregnancies, fetal anomalies diagnosed prenatally	**Screen:** 1 h, 50 g OGCT **Diagnostic:** 3 h, 100 g OGTT (CC, 1982; NDDG, 1979)	**G1:** Consultation with a dietitian; 2 weeks of diet restriction; fasting glucose level >105mg/dL, patient referred to endocrinologist, received glucose monitoring device, and began insulin treatment (n=489) **G2:** Did not receive further medical control (n=385)	Maternal hypertension, cesarean delivery, maternal birth trauma (postpartum hemorrhage), macrosomia, shoulder dystocia, mortality (intrauterine fetal demise), preterm labour, APGAR scores	NOS = 7 (good)

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
Crowther, 2005 Gillman, 2010 (4-5 year outcomes for children) Moss, 2007 (economic analysis) RCT, multi-center Sept 1993 to June 2003 Australia	1,000 **G1:** 30.9 ± 5.4 **G2:** 30.1 ± 5.5 **G1:** 26.8 (23.3–31.2) **G2:** 26.0 (22.9–30.9) **G1:** 4.8 ± 0.7 mmol/L **G2:** 4.8 ± 0.6 mmol/L **G1**: White: 356 Asian: 92 Other: 42 **G2**: White: 396 Asian: 72 Other: 42	**Inclusion:** Singleton or twin pregnancy; 16–30 wks gestation; prenatal clinic attendance; ≥1 risk factors for GDM on selective screen (WHO) or positive 50 g GCT and 75 g OGTT at 24–34 wks **Exclusion:** More severe glucose impairment; Hx of GDM; active chronic systemic disease	**Screen:** 50 g GCT (24–34 wks with 1h cutoff by WHO criteria, 1985) From 1998 onward any glucose level above normal classified as GDM (glucose level 1 h after GCT of at least 7.8 mmol/L) **Diagnostic:** 75 g OGTT (24–34 wks at fasting and 2-h) assessed by WHO criteria, 1985 From 1998 onward any glucose level above normal classified as GDM (venous plasma glucose level less than 6.1 – 7.0 mmol/L after overnight fast and 7.0–11.0 mmol/L at 2 h)	**G1:** Ongoing care; dietary advice; blood glucose monitoring; pre-prandial blood glucose target 5.5 mmol/L; 2 h 7.0 mmol/L; BG target of under 8.0 mmol/l was set at more than 35 weeks of pregnancy (n=490) **G2:** Replicated routine clinical care where GDM screening not available (n=510)	Induction of labor, caesarean delivery (elective & emergency), shoulder dystocia, hypoglycemia, hyperbilirubinemia, stillbirth or neonatal death, admission to NICU, birth weight, bone fracture/clavicular fracture, nerve palsy/brachial plexus injury, "Any serious prenatal complication", APGAR, LGA + SGA, 6 wk + 3 mo. Postpartum physical functioning, general health, vitality, emotional role, health state utility, anxiety, visits with healthcare professionals	RoB = Low (good)

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
Fassett, 2007 Cohort (with historical controls) Jan 2001 to June 2006 US	126 **G1:** 28.5 ± 5.8 **G2:** 29.2 ± 5.0 NR NR **G1**: White: 23 Black: 2 Hispanic: 39 Asian: 5 Other: 0 **G2**: White: 14 Black: 1 Hispanic: 35 Asian: 6 Other: 1	**Inclusion:** Women with ≥1 risk factors: prior GDM; prior macrosomia; first-degree relative with DM; prior stillbirth; prior malformation; 24–28 wks gestation; GDM Dx with CC criteria but not NDDG **Exclusion:** NR	**Screen:** 50 g GCT (24–28 wks with 1 h cutoff) **Diagnostic:** 100 g OGTT (24–28 wks at Fasting, 1 h, 2 h, and 3 h intervals) assessed by CC criteria	**G1:** Routine medical nutrition therapy by dietitian; formal diet (20–35 kcal/kg of prepregnancy body weight); BG daily self-monitoring, insulin as needed (n=69) **G2:** Historical controls before institution of routine medical nutrition therapy (n=57)	Caesarean delivery, unplanned caesarean delivery, weight gain, shoulder dystocia, admission to NICU, birth weight, neonatal metabolic complications, APGAR	NOS = 7 (good)

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
Garner, 1997 Malcolm, 2006 (7-11 yr f-up) RCT Sept 1991 to May 1994 Canada	300 **G1:** 30.7 ± 4.8 **G2:** 30.7 ± 4.6 NR **G1:** 180.0 ± 25.2 (10.0 ± 1.4 mmol/L) **G2:** 183.6 ± 32.4 mg/dL (10.2 ± 1.8 mmol/L) **G1**: NR **G2**: NR	**Inclusion:** Women with GDM diagnosed between 24–32 wks gestation; low-risk pregnancy **Exclusion:** Multiple gestation; maternal-fetal group incompatibility; known congenital anomaly; prior evidence of placenta previa or abruptio placentae; CHT; connective tissue disease; endocrine disorders; chronic hepatic disease; long-term medical therapy affecting glucose metabolism; imminent delivery	**Screen:** 75 g GCT (24–28 wks with 1 h cutoff by O'Sullivan criteria,1 h level of 144 mg/dL **Diagnostic:** 75 g OGTT (24–28 wks with Fasting ≥140 mg/dL, ≥11.1; 1 h, 2 h, and 3 h intervals) assessed by Hatem et al. criteria	**G1:** Strict glycemic control and tertiary level obstetric monitoring; dietary counseling, calorie-restricted diet, BG daily self-monitoring, insulin as needed (n=149) **G2:** Routine obstetric care (unrestricted healthy diet) (n=150)	Caesarean delivery, weight gain, hypoglycemia, hyperbilirubinemia, birth trauma, birth weight, child outcomes 7-11 yrs Normal 2 h GTT, at risk for overweight	RoB = High (poor)

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
Landon, 2009 RCT, multi-center Oct 2002 to Nov 2007 US	958 **G1:** 29.2 ± 5.7 **G2:** 28.9 ± 5.6 **G1:** 30.1 ± 5.0 **G2:** 30.2 ± 5.1 **G1**: fasting 86.6 ± 5.7 mg/dL (4.8 ± 0.3 mmol/L); 1 h 191.8 ± 21.9 mg/dL (10.7 ± 1.2 mmol/L); 2 h 173.7 ± 21.8 mg/dL (9.6 ±1.2 mmol/L); 3 h 137.3 ± 29.0 mg/dL (7.6 ±1.6 mmol/L) **G2:** fasting 86.3 ± 5.7 mg/dL (4.8 ± 0.3 mmol/L); 1 h 193.4 ± 19.3 mg/dL (10.7 ± 1.1 mmol/L); 2 h173.3 ± 19.6 mg/dL (9.6 ± 1.1 mmol/L); 3 h 134.1 ± 31.5 **G1**:White: 123 Black: 56 Hispanic: 281 Asian: 22 Other: 3 **G2**: White: 119 Black: 54 Hispanic: 265 Asian: 28 Other: 7	**Inclusion:** Women between 24 wks 0 days and 30 wks 6 days; OGCT values between 135 and 200 mg/dL or 7.5 and 11.1 mmol/L; OGTT fasting glucose <95 mg/dL and 2-3 timed measurements exceeded above thresholds at 1, 2, and 3 h. **Exclusion:** Abnormal result before 24 wks of gestation; preexisting diabetes; prior GDM; Hx of stillbirth; multifetal gestation; asthma; CHT; corticosteroid use; known fetal anomaly; likely preterm delivery	**Screen:** 50 g GCT (1-h cutoff) **Diagnostic:** 100 g OGTT (Fasting, 1 h, 2 h, and 3 h intervals) assessed by the 4th IWC criteria (1 h 180 mg/dL; 2 h 155 mg/dL; 3 h 140 mg/dL)	**G1:** Nutritional counseling and dietary therapy; daily BG self-monitoring; insulin as needed (n=485) **G2:** Usual perinatal care (n=473)	Induction of labor, caesarean delivery, preeclampsia, GHT, BMI at delivery, weight gain, shoulder dystocia, hypoglycemia, hyperbilirubine mia, elevated cord-blood c-peptide level, stillbirth or neonatal death, birth trauma, preterm delivery, admission to NICU, primary perinatal outcome, intravenous glucose Tx, respiratory distress syndrome, LGA, SGA, BMI at delivery	RoB = Unclear (fair)

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
Langer 2005 Cohort Jan 1990 to Sept 1999 US	2,775 **G1:** 29.1 ± 6 **G2:** 27.6 ± 6 **G1:** NR **G2:** NR **G1:** fasting 97 ± 16 mg/dL (5.4 mmol/L); 1 h 199 ± 28 mg/dL (11.1 mmol/L); 2 h 178 ± 30 (9.9 mmol/L); 3 h 136 ± 36 (7.5 mmol/L) **G2:** fasting 97 ± 15 mg/dL (5.4 mmol/L); 1 hr 199 ± 27 mg/dL (11.1 mmol/L); 2 hr 181 ± 36 mg/dL (10.1 mmol/L); 3 hr 141 ± 32 mg/dL 7.8 mmol/L) **G1**: White: 144 Black: 56 Hispanic: 910 **G2**: White: 61 Black: 17 Hispanic: 477	**Inclusion:** Singleton pregnancies; FPG < 140 mg/dL on OGTT; CASE CONTROL: GDM diagnosed > 37 wks; treated GDM and diabetic matched 2:1 obesity, parity, ethnicity, GA at delivery (within 5 days), and yr of delivery **Exclusion:** Pregestational DM; substance abusers; multifetal gestation; fetal anomalies	**Screen:** 50 g GCT (1 h >37 wks for G2; G1 underwent universal screening); Plasma glucose < 130 mg/dL **Diagnostic:** 100 g OGTT (>37 wks for G2; G1 underwent universal screening; Fasting, 1 h, 2 h, and 3 h intervals) assessed by CC criteria	**G1:** Diet alone or insulin and diet; formal diet with caloric restriction: 25 (overweight/obese) to 35 (normal weight) kcal/kg for actual pregnancy weight; 3 meals and 4 snacks; daily BG self-monitoring, insulin therapy if diet not successful in achieving glycemic control after 2 weeks (n=1,110) **G2:** Standard care until delivery (n=555)	Induction of labor, caesarean delivery, shoulder dystocia, hypoglycemia, stillbirth or neonatal death, birth weight, ponderal index, arterial cord <7.0, composite outcome, overall metabolic complications, erythrocytosis, respiratory complication, LGA, SGA	NOS = 9 (good)

Author, year Study Design Dates of study Country	Women Enrolled, *n* Maternal Age, *mean± SD (yr)* BMI, *mean ± SD; median IQR (kg/m^s)* *Glucose Levels, mean ± SD* *Race*	Inclusion/ Exclusion Criteria	Screening and Diagnostic Tests	Interventions	Outcomes Reported	Quality
Naylor, 1997 RCT Sept 1989 to Mar 1992 Canada	3,778 **G1:** 32.7(4.3) **G2**: 32.1 (4.4) **G1**: 24.2 (4.8) **G2:** 24.7(5.8) NR **G1:** White: 63 Black: 8 Asian: 27 Other: 45 **G2:** White: 67 Black: 2 Asian: 17 Other: 29	**Inclusion:** >24 yrs at time of delivery, no Hx of DM examined by physician before 24 wks gestation, delivery >28 wks ; **Exclusion:** NR	**Screen:** 50 g GCT (1 h - Plasma glucose < 130 mg/dL **Diagnostic:** 100 g OGTT assessed by NDDG criteria	**G1:** Known to have received treatment for GDM (n= 143) **G2:** Usual perinatal care (n= 115)	Preeclampsia, cesarean delivery, macrosomia, hypoglycemia, hyperbilirubinemia (phototherapy), fetal trauma, congenital anomalies, respiratory distress syndrome, maternal/fetal length of stay	NOS = 9 (good)

* ADA = American Diabetes Association; ADIPS = Australian Diabetes in Pregnancy Society; BMI = body mass index; CHT = chronic hypertension; d(s) = day(s); dL = deciliter; DM = diabetes mellitus; Dx = diagnosis/diagnostic; FPG = fasting plasma glucose; GCT = glucose tolerance test; GDM = gestational diabetes mellitus; GLT = glucose load test; g = grams; h = hour; IADPSG = International Association of Diabetes and Pregnancy Study Groups; JSOG = Japan Society of Obstetrics and Gynecology; mg = milligrams; NDDG = National Diabetes Data Group; NR = not reported; NOS = Newcastle-Ottawa Quality Assessment Scale; n = number; OGTT = oral glucose tolerance test; PP= postpartum; PCS = prospective cohort study; RCS = retrospective cohort study; RoB = Collaboration's tool for assessing risk of bias; SD = standard deviation; tx = treatment; wk(s) = week(s); WHO = World Health Organization; yr(s) = year(s)

Appendix E. List of Excluded Studies and Unobtained Studies

Excluded – Comparator (N=227)

1. Catalano PM, Avallone DA, Drago NM, et al. Reproducibility of the oral glucose tolerance test in pregnant women. Am J Obstet Gynecol 1993;169(4):874-81.

2. Schrader HM, Jovanovic-Peterson L, Bevier WC, et al. Fasting plasma glucose and glycosylated plasma protein at 24 to 28 weeks of gestation predict macrosomia in the general obstetric population. Am J Perinatol 1995;12(4):247-51.

3. Moses RG, Griffiths RD. Can a diagnosis of gestational diabetes be an advantage to the outcome of pregnancy? J Soc Gynecol Investig 1995;2(3):523-5.

4. Bassaw B, Ataullah I, Roopnarinesingh S, et al. Diabetes in pregnancy. Int J Gynaecol Obstet 1995;50(1):5-9.

5. de VM, Major CA, Morgan MA, et al. Postprandial versus preprandial blood glucose monitoring in women with gestational diabetes mellitus requiring insulin therapy. N Engl J Med 1995;333(19):1237-41.

6. Damm P, Kuhl C, Hornnes P, et al. A longitudinal study of plasma insulin and glucagon in women with previous gestational diabetes. Diabetes Care 1995;18(5):654-65.

7. Koukkou E, Taub N, Jackson P, et al. Difference in prevalence of gestational diabetes and perinatal outcome in an innercity multiethnic London population. Eur J Obstet Gynecol Reprod Biol 1995;59(2):153-7.

8. Gribble RK, Meier PR, Berg RL. Blood glucose limits in the diagnosis of impaired glucose tolerance during pregnancy. Relation to morbidity. Obstet Gynecol 1995;86(3):405-10.

9. al-Najashi SS. Control of gestational diabetes. Int J Gynaecol Obstet 1995;49(2):131-5.

10. Kjos SL, Peters RK, Xiang A, et al. Predicting future diabetes in Latino women with gestational diabetes. Utility of early postpartum glucose tolerance testing. Diabetes 1995;44(5):586-91.

11. Reece EA, Hagay Z, Gay LJ, et al. A randomized clinical trial of a fiber-enriched diabetic diet vs. the standard American Diabetes Association-recommended diet in the management of diabetes mellitus in pregnancy. J Matern Fetal Invest 1995;5(1):8-12.

12. Hughes PF, Agarwal M, Newman P, et al. An evaluation of fructosamine estimation in screening for gestational diabetes mellitus. Diabet Med 1995;12(8):708-12.

13. Agardh CD, Aberg A, Norden NE. Glucose levels and insulin secretion during a 75 g glucose challenge test in normal pregnancy. J Intern Med 1996;240(5):303-9.

14. Hod M, Rabinerson D, Kaplan B, et al. Perinatal complications following gestational diabetes mellitus how 'sweet' is ill? Acta Obstet Gynecol Scand 1996;75(9):809-15.

15. Nasrat HA, Ardawi MS, Abalkhail BA. The diagnosis of "pathological hyperglycaemia' in gestational diabetes in a high risk obstetric population. Diabetic Med 1996;13(10):861-7.

16. Di SN, Ronsisvalle E, Fulghesu AM, et al. Insulin plasma levels in pregnant patients with impaired glucose tolerance: relationship with pregnancy outcome. Gynecol Obstet Invest 1996;42(1):16-20.

17. Tranquilli AL, Pizzichini L, Cingolani F, et al. Prediction of the need for insulin therapy in pregnant women with impaired gestational glucose tolerance (IGGT). Clin Exp Obstet Gynecol 1996;23(2):79-82.

18. Moses RG. The recurrence rate of gestational diabetes in subsequent pregnancies. Diabetes Care 1996;19(12):1348-50.

19. Di CG, Benzi L, Casadidio I, et al. Screening of gestational diabetes in Tuscany: results in 2000 cases. Ann Ist Super Sanita 1997;33(3):389-91.

20. Fedele D, Lapolla A. A protocol of screening of gestational diabetes mellitus. Ann Ist Super Sanita 1997;33(3):383-7.

21. Bienstock JL, Blakemore KJ, Wang E, et al. Managed care does not lower costs but may result in poorer outcomes for patients with gestational diabetes. Am J Obstet Gynecol 1997;177(5):1035-7.

22. Helton MR, Arndt J, Kebede M, et al. Do low-risk prenatal patients really need a screening glucose challenge test? J Fam Pract 1997;44(6):556-61.

23. Avery MD, Leon AS, Kopher RA. Effects of a partially home-based exercise program for women with gestational diabetes. Obstet Gynecol 1997;89(1):10-5.

24. al-Najashi SS. Diagnosis of gestational diabetes mellitus. A study of 315 pregnant women at King Fahd Hospital of the University, Al-Khobar, Saudi Arabia. Bahrain Med Bull 1997;19(4):104-7.

25. Persson B, Edwall L, Hanson U, et al. Insulin sensitivity and insulin response in women with gestational diabetes mellitus. Horm Metab Res 1997;29(8):393-7.

26. Berria R, Murgia C, Serri F, Pilia I. GDM: screening, diagnosis and management. 1997.

27. Casey BM, Lucas MJ, McIntire DD, et al. Pregnancy outcomes in women with gestational diabetes compared with the general obstetric population. Obstet Gynecol 1997;90(6):869-73.

28. Lurie S, Levy R, Weiss R, et al. Low values on 50 gram glucose challenge test or oral 100 gram glucose tolerance test are associated with good perinatal outcome. J Obstet Gynaecol 1998;18(5):451-4.

29. Ramachandran A, Snehalatha C, Clementina M, et al. Foetal outcome in gestational diabetes in south Indians. Diabetes Res Clin Pract 1998;41(3):185-9.

30. Whitaker RC, Pepe MS, Seidel KD, et al. Gestational diabetes and the risk of offspring obesity. Pediatrics 1998;101(2):e9.

31. Kitzmiller JL, Elixhauser A, Carr S, et al. Assessment of costs and benefits of management of gestational diabetes mellitus. Diabetes Care 1998;21(Suppl 2):B123-B130.

32. Schafer-Graf UM, Dupak J, Vogel M, et al. Hyperinsulinism, neonatal obesity and placental immaturity in infants born to women with one abnormal glucose tolerance test value. J Perinat Med 1998;26(1):27-36.

33. Major CA, Henry MJ, de VM, et al. The effects of carbohydrate restriction in patients with diet-controlled gestational diabetes. Obstet Gynecol 1998;91(4):600-4.

34. Lao TT, Lee CP. Gestational 'impaired glucose tolerance': should the cut-off be raised to 9 mmol l(-1)? Diabetic Med 1998;15(1):25-9.

35. Moses RG, Moses J, Davis WS. Gestational diabetes: Do lean young Caucasian women need to be tested? Diabetes Care 1998;21(11):1803-6.

36. Moses RG, Moses M, Russell KG, et al. The 75-g glucose tolerance test in pregnancy - A reference range determined on a low-risk population and related to selected pregnancy outcomes. Diabetes Care 1998;21(11):1807-11.

37. Lemen PM, Wigton TR, Miller-McCarthey AJ, et al. Screening for gestational diabetes mellitus in adolescent pregnancies. Am J Obstet Gynecol 1998;178(6):1251-4.

38. Lauszus FF, Paludan J, Klebe JG. Birthweight in women with potential gestational diabetes mellitus - An effect of obesity rather than glucose intolerance? Acta Obstet Gynecol Scand 1999;78(6):520-5.

39. Roncaglia N, Bellini P, Arreghini A, et al. Gestational diabetes mellitus: intensive versus mild treatment. Clin Exp Obstet Gynecol 1999;26(2):95-7.

40. McFarland MB, Langer O, Conway DL, et al. Dietary therapy for gestational diabetes: how long is long enough? Obstet Gynecol 1999;93(6):978-82.

41. Kvetny J, Poulsen HF, Damgaard DW. Results from screening for gestational diabetes mellitus in a Danish county. Dan Med Bull 1999;46(1):57-9.

42. Shivvers SA, Lucas MJ. Gestational diabetes - Is a 50-g screening result >= 200 mg/dL diagnostic? J Reprod Med 1999;44(8):685-8.

43. Jovanovic L, Gutierrez M, Peterson CM. Chromium supplementation for women with gestational diabetes mellitus. J Trace Elem Exp Med 1999;12(2):91-7.

44. Vohr BR, McGarvey ST, Tucker R. Effects of maternal gestational diabetes on offspring adiposity at 4-7 years of age. Diabetes Care 1999;22(8):1284-91.

45. Mirghani OA, Saeed OK. A simplified management of diabetic pregnant woman. Saudi Med J 2000;21(4):335-9.

46. Rae A, Bond D, Evans S, et al. A randomised controlled trial of dietary energy restriction in the

management of obese women with gestational diabetes. Aust N Z J Obstet Gynaecol 2000;40(4):416-22.

47. Langer O, Conway DL, Berkus MD, et al. A comparison of glyburide and insulin in women with gestational diabetes mellitus. N Engl J Med 2000;343(16):1134-8.

48. Rudge MV, Calderon IM, Ramos MD, et al. Perinatal outcome of pregnancies complicated by diabetes and by maternal daily hyperglycemia not related to diabetes. A retrospective 10-year analysis. Gynecol Obstet Invest 2000;50(2):108-12.

49. Rigano S, Ferrazzi E, Radaelli T, et al. Sonographic measurements of subcutaneous fetal fat in pregnancies complicated by gestational diabetes and in normal pregnancies. Croat Med J 2000;41(3):240-4.

50. Bancroft K, Tuffnell DJ, Mason GC, et al. A randomised controlled pilot study of the management of gestational impaired glucose tolerance. BJOG 2000;107(8):959-63.

51. Lao TT, Ho LF. Impaired glucose tolerance and pregnancy outcome in Chinese women with high body mass index. Hum Reprod 2000;15(8):1826-9.

52. Simpson RW, Kast SJ. Management of gestational diabetes with a conservative insulin protocol. The Medical Journal Of Australia 2000;172(11):537-40.

53. Hearty RT, Traub AI, Hadden DR. Screening for hyperglycaemia in pregnancy: analysis of two screening protocols and review of current methods. Ulster Med J 2000;69(1):35-43.

54. Crowe SM, Mastrobattista JM, Monga M. Oral glucose tolerance test and the preparatory diet. Am J Obstet Gynecol 2000;182(5):1052-4.

55. Wong L, Tan AS. The glucose challenge test for screening gestational diabetes in pregnant women with no risk factors. Singapore Med J 2001;42(11):517-21.

56. Lauszus FF, Rasmussen OW, Henriksen JE, et al. Effect of a high monounsaturated fatty acid diet on blood pressure and glucose metabolism in women with gestational diabetes mellitus. Eur J Clin Nutr 2001;55(6):436-43.

57. Jensen DM, Damm P, Sorensen B, et al. Clinical impact of mild carbohydrate intolerance in pregnancy: a study of 2904 nondiabetic Danish women with risk factors for gestational diabetes mellitus. Am J Obstet Gynecol 2001;185(2):413-9.

58. MacNeill S, Dodds L, Hamilton DC, et al. Rates and risk factors for recurrence of gestational diabetes. Diabetes Care 2001;24(4):659-62.

59. Ray JG, Vermeulen MJ, Shapiro JL, et al. Maternal and neonatal outcomes in pregestational and gestational diabetes mellitus, and the influence of maternal obesity and weight gain: the DEPOSIT study. Diabetes Endocrine Pregnancy Outcome Study in Toronto. QJM 2001;94(7):347-56.

60. Xiong X, Saunders LD, Wang FL, et al. Gestational diabetes mellitus: prevalence, risk factors, maternal and infant outcomes. Int J Gynaecol Obstet 2001;75(3):221-8.

61. Baliutaviciene D, Petrenko V, Zalinkevicius R. Selective or universal diagnostic testing for gestational diabetes mellitus. Int J Gynaecol Obstet 2002;78(3):207-11.

62. Gezer A, Esen F, Mutlu H, et al. Prognosis of patients with positive screening but negative diagnostic test for gestational diabetes. Arch Gynecol Obstet 2002;266(4):201-4.

63. Di CG, Volpe L, Casadidio I, et al. Universal screening and intensive metabolic management of gestational diabetes: cost-effectiveness in Italy. Acta Diabetol 2002;39(2):69-73.

64. Homko CJ, Sivan E, Reece EA. The impact of self-monitoring of blood glucose on self-efficacy and pregnancy outcomes in women with diet-controlled gestational diabetes. Diabetes Educ 2002;28(3):435-43.

65. Rumbold AR, Crowther CA. Women's experiences of being screened for gestational diabetes mellitus. Aust N Z J Obstet Gynaecol 2002;42(2):131-7.

66. Pullen F, Grenfell A. The diagnosis of gestational diabetes in a multiethnic population: Which diagnostic criteria should be used with respect to maternal outcome? Pract Diabetes Int 2002;29(9):279-82.

67. Pettitt DJ, Ospina P, Jovanovic L. Comparison of an insulin analog, insulin aspart, and regular human insulin with no insulin in gestational diabetes mellitus. Diabetologia 2002;45(Suppl 2):A254-A255.

68. Chen R, Yogev Y, Ben-Haroush A, et al. Continuous glucose monitoring for the evaluation and improved control of gestational diabetes mellitus. J Matern Fetal Neonatal Med 2003;14(4):256-60.

69. Mecacci F, Carignani L, Cioni R, et al. Maternal metabolic control and perinatal outcome in women with gestational diabetes treated with regular or lispro insulin: comparison with non-diabetic pregnant women. Eur J Obstet Gynecol Reprod Biol 2003;111(1):19-24.

70. Gruendhammer M, Brezinka C, Lechleitner M. The number of abnormal plasma glucose values in the oral glucose tolerance test and the feto-maternal outcome of pregnancy. Eur J Obstet Gynecol Reprod Biol 2003;108(2):131-6.

71. Kulkarni M, Jones KD, Newbold S. Screening for gestational diabetes: a retrospective audit. J Obstet Gynaecol 2003;23(2):160-2.

72. Daniells S, Grenyer BFS, Davis WS, et al. Gestational diabetes mellitus: Is a diagnosis associated with an increase in maternal anxiety and stress in the short and intermediate term? Diabetes Care 2003;26(2):385-9.

73. Saldana TM, Siega-Riz AM, Adair LS, et al. The association between impaired glucose tolerance and birth weight among black and white women in central North Carolina. Diabetes Care 2003;26(3):656-61.

74. Sunsaneevithayakul P, Ruangvutilert P, Sutanthavibul A, et al. Effect of 3-day intensive dietary therapy during admission in women after diagnosis of gestational diabetes mellitus. J Med Assoc Thai 2004;87(9):1022-8.

75. Ertunc D, Tok E, Dilek U, et al. The effect of carbohydrate intolerance on neonatal birth weight in pregnant women without gestational diabetes mellitus. Ann Saudi Med 2004;24(4):280-3.

76. Bonomo M, Cetin I, Pisoni MP, et al. Flexible treatment of gestational diabetes modulated on ultrasound evaluation of intrauterine growth: a controlled randomized clinical trial. Diabetes Metab 2004;30(3):237-44.

77. Maser RE, Lenhard MJ, Henderson BC, et al. Detection of subsequent episodes of gestational diabetes mellitus: a need for specific guidelines. J Diabetes Complications 2004;18(2):86-90.

78. Conway DL, Gonzales O, Skiver D. Use of glyburide for the treatment of gestational diabetes: the San Antonio experience. J Matern Fetal Neonatal Med 2004;15(1):51-5.

79. Brankston GN, Mitchell BF, Ryan EA, et al. Resistance exercise decreases the need for insulin in overweight women with gestational diabetes mellitus. Am J Obstet Gynecol 2004;190(1):188-93.

80. Schaefer-Graf UM, Kjos SL, Fauzan OH, et al. A randomized trial evaluating a predominantly fetal growth-based strategy to guide management of gestational diabetes in Caucasian women. Diabetes Care 2004;27(2):297-302.

81. Lauenborg J, Hansen T, Jensen DM, et al. Increasing Incidence of Diabetes after Gestational Diabetes: A long-term follow-up in a Danish population. Diabetes Care 2004;27(5):1194-9.

82. Kim C, Brawarsky P, Jackson RA, et al. Changes in health status experienced by women with gestational diabetes and pregnancy-induced hypertensive disorders. J Womens Health 2005;14(8):729-36.

83. Bertini AM, Silva JC, Taborda W, et al. Perinatal outcomes and the use of oral hypoglycemic agents. J Perinat Med 2005;33(6):519-23.

84. Sharpe PB, Chan A, Haan EA, et al. Maternal diabetes and congenital anomalies in South Australia 1986-2000: a population-based cohort study. Birth Defects Res Part A Clin Mol Teratol 2005;73(9):605-11.

85. Leipold H, Worda C, Gruber CJ, et al. Large-for-gestational-age newborns in women with insulin-treated gestational diabetes under strict metabolic control. Wien Klin Wochenschr 2005;117(15-16):521-5.

86. Chan BC, Lao TT. Gestational diabetes mellitus in women in the fourth decade--is treatment worthwhile? Gynecol Obstet Invest 2005;60(2):112-6.

87. Jacobson GF, Ramos GA, Ching JY, et al. Comparison of glyburide and insulin for the management of gestational diabetes in a large managed care organization. Am J Obstet Gynecol 2005;193(1):118-24.

88. Barahona MJ, Sucunza N, Garcia-Patterson A, et al. Period of gestational diabetes mellitus diagnosis and maternal and fetal morbidity. Acta Obstet Gynecol Scand 2005;84(7):622-7.

89. Yogev Y, Langer O, Xenakis EM, et al. The association between glucose challenge test, obesity and pregnancy outcome in 6390 non-diabetic women. J Matern Fetal Neonatal Med 2005;17(1):29-34.

90. Weisz B, Shrim A, Homko CJ, et al. One hour versus two hours postprandial glucose measurement in gestational diabetes: a prospective study. J Perinatol 2005;25(4):241-4.

91. Ho LF, Benzie IF, Lao TT. Relationship between caloric intake and pregnancy outcome in diet-treated gestational diabetes mellitus. Nurs Health Sci 2005;7(1):15-20.

92. Langer O, Yogev Y, Xenakis EM, et al. Insulin and glyburide therapy: dosage, severity level of gestational diabetes, and pregnancy outcome. Am J Obstet Gynecol 2005;192(1):134-9.

93. Keshavarz M, Cheung NW, Babaee GR, et al. Gestational diabetes in Iran: incidence, risk factors and pregnancy outcomes. Diabetes Res Clin Pract 2005;69(3):279-86.

94. Ezimokhai M, Joseph A, Bradley-Watson P. Audit of pregnancies complicated by diabetes from one center five years apart with selective versus universal screening. Ann N Y Acad Sci 2006;1084:132-40.

95. Li K, Yang HX. Value of fructosamine measurement in pregnant women with abnormal glucose tolerance. Chin Med J 2006;119(22):1861-5.

96. Thanasuan S, Borriboonhirunsarn D. Incidence of gestational diabetes mellitus among pregnant women with one abnormal value of oral glucose tolerance test. J Med Assoc Thai 2006;89(8):1109-14.

97. Rochon M, Rand L, Roth L, et al. Glyburide for the management of gestational diabetes: risk factors predictive of failure and associated pregnancy outcomes. Am J Obstet Gynecol 2006;195(4):1090-4.

98. Reader D, Splett P, Gunderson EP, et al. Impact of gestational diabetes mellitus nutrition practice guidelines implemented by registered dietitians on pregnancy outcomes. J Am Diet Assoc 2006;106(9):1426-33.

99. Shefali AK, Kavitha M, Deepa R, et al. Pregnancy outcomes in pre-gestational and gestational diabetic women in comparison to non-diabetic women--A prospective study in Asian Indian mothers (CURES-35). J Assoc Physicians India 2006;54:613-8.

100. Chandna A, Zuberi LM, Munim S. Threshold values for the glucose challenge test in pregnancy. Int J Gynaecol Obstet 2006;94(2):119-20.

101. Kahn BF, Davies JK, Lynch AM, et al. Predictors of glyburide failure in the treatment of gestational diabetes. Obstet Gynecol 2006;107(6):1303-9.

102. Dudhbhai M, Lim L, Bombard A, et al. Characteristics of patients with abnormal glucose challenge test and normal oral glucose tolerance test results: comparison with normal and gestational diabetic patients. Am J Obstet Gynecol 2006;194(5):e42-e45.

103. McLaughlin GB, Cheng YW, Caughey AB. Women with one elevated 3-hour glucose tolerance test value: are they at risk for adverse perinatal outcomes? Am J Obstet Gynecol 2006;194(5):e16-e19.

104. Sunsaneevithayakul P, Kanokpongsakdi S, Sutanthavibul A, et al. Result of ambulatory diet therapy in gestational diabetes mellitus. J Med Assoc Thai 2006;89(1):8-12.

105. D'Anna R, Baviera G, De VA, et al. C-reactive protein as an early predictor of gestational diabetes mellitus. J Reprod Med 2006;51(1):55-8.

106. Nordin NM, Wei JW, Naing NN, et al. Comparison of maternal-fetal outcomes in gestational diabetes and lesser degrees of glucose intolerance. J Obstet Gynaecol Res 2006;32(1):107-14.

107. Weijers RN, Bekedam DJ, Goldschmidt HM, et al. The clinical usefulness of glucose tolerance testing in gestational diabetes to predict early postpartum diabetes mellitus. Clin Chem Lab Med 2006;44(1):99-104.

108. Cosson E, Benchimol M, Carbillon L, et al. Universal rather than selective screening for gestational diabetes mellitus may improve fetal outcomes. Diabetes Metab 2006;32(2):140-6.

109. Moore LE, Briery CM, Clokey D, et al. Metformin and insulin in the management of gestational diabetes mellitus: preliminary results of a comparison. J Reprod Med 2007;52(11):1011-5.

110. Most O, Langer O. GDM women in good glycemic control: which meal-related measure enhances fetal well-being? J Perinat Med 2007;35(6):481-5.

111. Cypryk K, Kaminska P, Kosinski M, et al. A comparison of the effectiveness, tolerability and safety of high and low carbohydrate diets in women with gestational diabetes. Endokrynol Pol 2007;58(4):314-9.

112. Cheng YW, McLaughlin GB, Esakoff TF, et al. Glucose challenge test: screening threshold for gestational diabetes mellitus and associated outcomes. J Matern Fetal Neonatal Med 2007;20(12):903-8.

113. Krishnaveni GV, Hill JC, Veena SR, et al. Gestational diabetes and the incidence of diabetes in the 5 years following the index pregnancy in South Indian women. Diabetes Res Clin Pract 2007;78(3):398-404.

114. Virally M, Laloi-Michelin M, Meas T, et al. Occurrence of gestational diabetes mellitus, maternal and fetal outcomes beyond the 28th week of gestation in women at high risk of gestational diabetes. A prospective study. Diabetes Metab 2007;33(4):290-5.

115. Pettitt DJ, Ospina P, Howard C, et al. Efficacy, safety and lack of immunogenicity of insulin aspart compared with regular human insulin for women with gestational diabetes mellitus. Diabetic Med 2007;24(10):1129-35.

116. Yogev Y, Langer O. Spontaneous preterm delivery and gestational diabetes: the impact of glycemic control. Arch Gynecol Obstet 2007;276(4):361-5.

117. Todorova K, Palaveev O, Petkova VB, et al. A pharmacoeconomical model for choice of a treatment for pregnant women with gestational diabetes. Acta Diabetol 2007;44(3):144-8.

118. Dodd JM, Crowther CA, Antoniou G, et al. Screening for gestational diabetes: the effect of varying blood glucose definitions in the prediction of adverse maternal and infant health outcomes. Aust N Z J Obstet Gynaecol 2007;47(4):307-12.

119. Rowan JA, MiG I. A trial in progress: gestational diabetes. Treatment with metformin compared with insulin (the Metformin in Gestational Diabetes [MiG] trial).[Erratum appears in Diabetes Care. 2007 Dec;30(12):3154]. Diabetes Care 2007;30(Suppl 2):S214-S219.

120. Di CG, Seghieri G, Lencioni C, et al. Normal glucose tolerance and gestational diabetes mellitus: what is in between? Diabetes Care 2007;30(7):1783-8.

121. Homko CJ, Santamore WP, Whiteman V, et al. Use of an internet-based telemedicine system to manage underserved women with gestational diabetes mellitus. Diabetes Technol Ther 2007;9(3):297-306.

122. Artal R, Catanzaro RB, Gavard JA, et al. A lifestyle intervention of weight-gain restriction: diet and exercise in obese women with gestational diabetes mellitus. Appl Physiol Nutr Metab 2007;32(3):596-601.

123. Kestila KK, Ekblad UU, Ronnemaa T. Continuous glucose monitoring versus self-monitoring of blood glucose in the treatment of gestational diabetes mellitus. Diabetes Res Clin Pract 2007;77(2):174-9.

124. Ramos GA, Jacobson GF, Kirby RS, et al. Comparison of glyburide and insulin for the management of gestational diabetics with markedly elevated oral glucose challenge test and fasting hyperglycemia. J Perinatol 2007;27(5):262-7.

125. Simmons D. Relationship between maternal glycaemia and birth weight in glucose-tolerant women from different ethnic groups in New Zealand. Diabetic Med 2007;24(3):240-4.

126. Lee AJ, Hiscock RJ, Wein P, et al. Gestational diabetes mellitus: clinical predictors and long-term risk of developing type 2 diabetes: a retrospective cohort study using survival analysis. Diabetes Care 2007;30(4):878-83.

127. Price N, Bartlett C, Gillmer M. Use of insulin glargine during pregnancy: a case-control pilot study. BJOG 2007;114(4):453-7.

128. Berg M, Adlerberth A, Sultan B, et al. Early random capillary glucose level screening and multidisciplinary antenatal teamwork to improve outcome in gestational diabetes mellitus. Acta Obstet Gynecol Scand 2007;86(3):283-90.

129. Sinclair BA, Rowan JA, Hainsworth OT. Macrosomic infants are not all equal. Aust N Z J Obstet Gynaecol 2007;47(2):101-5.

130. Koklanaris N, Bonnano C, Seubert D, et al. Does raising the glucose challenge test threshold impact birthweight in Asian gravidas? J Perinat Med 2007;35(2):100-3.

131. Anjalakshi C, Balaji V, Balaji MS, et al. A prospective study comparing insulin and glibenclamide in gestational diabetes mellitus in Asian Indian women. Diabetes Res Clin Pract 2007;76(3):474-5.

132. Jasbinder K, Shivani J, Anju H, et al. Pregnancy outcome in gestational diabetes mellitus: continued risk related to FBS levels. Diabetes Res Clin Pract 2007;78(2):302-3.

133. Anderberg E, Kallen K, Berntorp K, et al. A simplified oral glucose tolerance test in pregnancy: compliance and results. Acta Obstet Gynecol Scand 2007;86(12):1432-6.

134. Ozcimen EE, Uckuyu A, Ciftci FC, et al. Diagnosis of gestational diabetes mellitus by use of the homeostasis model assessment-insulin resistance index in the first trimester. Gynecol Endocrinol 2008;24(4):224-9.

135. Tam WH, Ma RC, Yang X, et al. Glucose intolerance and cardiometabolic risk in children

exposed to maternal gestational diabetes mellitus in utero. Pediatrics 2008;122(6):1229-34.

136. Akinci B, Celtik A, Yener S, et al. Is fasting glucose level during oral glucose tolerance test an indicator of the insulin need in gestational diabetes? Diabetes Res Clin Pract 2008;82(2):219-25.

137. Holt RI, Clarke P, Parry EC, et al. The effectiveness of glibenclamide in women with gestational diabetes. Diabetes Obes Metab 2008;10(10):906-11.

138. Snapp CA, Donaldson SK. Gestational diabetes mellitus: physical exercise and health outcomes. Biol Res Nurs 2008;10(2):145-55.

139. Ju H, Rumbold AR, Willson KJ, et al. Borderline gestational diabetes mellitus and pregnancy outcomes. BMC Pregnancy Childbirth 2008;8:31.

140. Suhonen L, Hiilesmaa V, Kaaja R, et al. Detection of pregnancies with high risk of fetal macrosomia among women with gestational diabetes mellitus. Acta Obstet Gynecol Scand 2008;87(9):940-5.

141. Mendelson SG, Neese-Smith D, Koniak-Griffin D, et al. A community-based parish nurse intervention program for Mexican American women with gestational diabetes. J Obstet Gynecol & Neonatal Nurs 2008;37(4):415-25.

142. Gumus II, Turhan NO. Are patients with positive screening but negative diagnostic test for gestational diabetes under risk for adverse pregnancy outcome? J Obstet Gynaecol Res 2008;34(3):359-63.

143. Lee H, Jang HC, Park HK, et al. Prevalence of type 2 diabetes among women with a previous history of gestational diabetes mellitus. Diabetes Res Clin Pract 2008;81(1):124-9.

144. Rowan JA, Hague WM, Gao W, et al. Metformin versus insulin for the treatment of gestational diabetes.[Erratum appears in N Engl J Med. 2008 Jul 3;359(1):106]. N Engl J Med 2008;358(19):2003-15.

145. Grotegut CA, Tatineni H, Dandolu V, et al. Obstetric outcomes with a false-positive one-hour glucose challenge test by the Carpenter-Coustan criteria. J Matern Fetal Neonatal Med 2008;21(5):315-20.

146. Negrato CA, Jovanovic L, Tambascia MA, et al. Mild gestational hyperglycaemia as a risk factor for metabolic syndrome in pregnancy and adverse perinatal outcomes. Diabetes Metab Res Rev 2008;24(4):324-30.

147. Savona-Ventura C, Chircop M. Significant thresholds for the 75-g oral glucose tolerance test in pregnancy. J Diabetes Complications 2008;22(3):178-80.

148. Lapolla A, Dalfra MG, Mello G, et al. Early detection of insulin sensitivity and beta-cell function with simple tests indicates future derangements in late pregnancy. J Clin Endocrinol Metab 2008;93(3):876-80.

149. Hawkins JS, Lo JY, Casey BM, et al. Diet-treated gestational diabetes mellitus: comparison of early vs routine diagnosis. Am J Obstet Gynecol 2008;198(3):287-e1-6.

150. Keely EJ, Malcolm JC, Hadjiyannakis S, et al. Prevalence of metabolic markers of insulin resistance in offspring of gestational diabetes pregnancies. Pediatr Diabetes 2008;9(1):53-9.

151. Elnour AA, El Mugammar IT, Jaber T, et al. Pharmaceutical care of patients with gestational diabetes mellitus. Journal of Evaluation in Clinical Practice 2008;14(1):131-40.

152. Elnour AA, McElnay JC. Antenatal oral glucose-tolerance test values and pregnancy outcomes. Int J Pharm Pract 2008;16(3):189-97.

153. Carr DB, Newton KM, Utzschneider KM, et al. Modestly elevated glucose levels during pregnancy are associated with a higher risk of future diabetes among women without gestational diabetes mellitus. Diabetes Care 2008;31(5):1037-9.

154. Jensen DM, Korsholm L, Ovesen P, et al. Adverse pregnancy outcome in women with mild glucose intolerance: is there a clinically meaningful threshold value for glucose? Acta Obstet Gynecol Scand 2008;87(1):59-62.

155. Seshiah V, Balaji V, Balaji MS, et al. Prevalence of gestational diabetes mellitus in South India (Tamil Nadu)--a community based study. J Assoc Physicians India 2008;56:329-33.

156. Rai L, Meenakshi D, Kamath A. Metformin--a convenient alternative to insulin for Indian women with diabetes in pregnancy. Indian J Med Sci 2009;63(11):491-7.

157. Hebert MF, Ma X, Naraharisetti SB, et al. Are we optimizing gestational diabetes treatment with glyburide? The pharmacologic basis for better clinical practice. Clin Pharmacol Ther 2009;85(6):607-14.

158. Negrato CA, Jovanovic L, Tambascia MA, et al. Association between insulin resistance, glucose

intolerance, and hypertension in pregnancy. Metab Syndr Relat Disord 2009;7(1):53-9.

159. Segregur J, Bukovic D, Milinovic D, et al. Fetal macrosomia in pregnant women with gestational diabetes. Coll Antropol 2009;33(4):1121-7.

160. Pedula KL, Hillier TA, Schmidt MM, et al. Ethnic differences in gestational oral glucose screening in a large US population. Ethn Dis 2009;19(4):414-9.

161. Persson M, Winkvist A, Mogren I. Surprisingly low compliance to local guidelines for risk factor based screening for gestational diabetes mellitus - A population-based study. BMC Pregnancy Childbirth 2009;9:53.

162. Perichart-Perera O, Balas-Nakash M, Parra-Covarrubias A, et al. A medical nutrition therapy program improves perinatal outcomes in Mexican pregnant women with gestational diabetes and type 2 diabetes mellitus. Diabetes Educ 2009;35(6):1004-13.

163. Wroblewska-Seniuk K, Wender-Ozegowska E, Szczapa J. Long-term effects of diabetes during pregnancy on the offspring. Pediatr Diabetes2009;10(7):432-40.

164. Balani J, Hyer SL, Rodin DA, et al. Pregnancy outcomes in women with gestational diabetes treated with metformin or insulin: a case-control study. Diabetic Med 2009;26(8):798-802.

165. Retnakaran R, Shah BR. Abnormal screening glucose challenge test in pregnancy and future risk of diabetes in young women. Diabetic Med 2009;26(5):474-7.

166. Lapolla A, Dalfra MG, Bonomo M, et al. Gestational diabetes mellitus in Italy: a multicenter study. Eur J Obstet Gynecol Reprod Biol 2009;145(2):149-53.

167. Madarasz E, Tamas G, Tabak AG, et al. Carbohydrate metabolism and cardiovascular risk factors 4 years after a pregnancy complicated by gestational diabetes. Diabetes Res Clin Pract 2009;85(2):197-202.

168. Lin CH, Wen SF, Wu YH, et al. Using the 100-g oral glucose tolerance test to predict fetal and maternal outcomes in women with gestational diabetes mellitus. Chang Gung Medical Journal 2009;32(3):283-9.

169. Esakoff TF, Cheng YW, Sparks TN, et al. The association between birthweight 4000 g or greater and perinatal outcomes in patients with and without gestational diabetes mellitus. Am J Obstet Gynecol 2009;200(6):672-4.

170. Moses RG, Barker M, Winter M, et al. Can a low-glycemic index diet reduce the need for insulin in gestational diabetes mellitus? A randomized trial. Diabetes Care 2009;32(6):996-1000.

171. Barrett HL, Morris J, McElduff A. Watchful waiting: a management protocol for maternal glycaemia in the peripartum period. Aust N Z J Obstet Gynaecol 2009;49(2):162-7.

172. Lain KY, Garabedian MJ, Daftary A, et al. Neonatal adiposity following maternal treatment of gestational diabetes with glyburide compared with insulin. Am J Obstet Gynecol 2009;200(5):501-6.

173. Karmon A, Levy A, Holcberg G, et al. Decreased perinatal mortality among women with diet-controlled gestational diabetes mellitus. Int J Gynaecol Obstet 2009;104(3):199-202.

174. Most OL, Kim JH, Arslan AA, et al. Maternal and neonatal outcomes in early glucose tolerance testing in an obstetric population in New York city. J Perinat Med 2009;37(2):114-7.

175. Kucuk M, Doymaz F. Placental weight and placental weight-to-birth weight ratio are increased in diet- and exercise-treated gestational diabetes mellitus subjects but not in subjects with one abnormal value on 100-g oral glucose tolerance test. J Diabetes Complications 2009;23(1):25-31.

176. Pawelec M, Karmowski A, Krzemieniewska J, et al. The clinical and financial effects of replacing the 1 h 50 g screening test for gestational diabetes mellitus by the stick method. Adv Clin Exp Med 2009;18(6):601-7.

177. Hossein-nezhad A, Mirzaei K, Maghbooli Z, et al. Maternal glycemic status in GDM patients after delivery. Iran J Diabetes Lipid Disord 2009;8(1):95-104.

178. Clausen TD, Mathiesen ER, Hansen T, et al. Overweight and the metabolic syndrome in adult offspring of women with diet-treated gestational diabetes mellitus or type 1 diabetes. J Clin Endocrinol Metab 2009;94(7):2464-70.

179. Riskin-Mashiah S, Younes G, Damti A, et al. First-Trimester Fasting Hyperglycemia and Adverse Pregnancy Outcomes. Diabetes Care 2009;32(9):1639-43.

180. Herring SJ, Oken E, Rifas-Shiman SL, et al. Weight gain in pregnancy and risk of maternal

hyperglycemia. Am J Obstet Gynecol 2009;201(1):61-7.

181. Chen YZ, Quick WW, Yang WY, et al. Cost of Gestational Diabetes Mellitus in the United States in 2007. Population Health Management 2009;12(3):165-74.

182. Hedderson MM, Darbinian JA, Ferrara A. Disparities in the risk of gestational diabetes by race-ethnicity and country of birth. Paediatr Perinat Epidemiol 2010;24(5):441-8.

183. de Barros MC, Lopes MA, Francisco RP, et al. Resistance exercise and glycemic control in women with gestational diabetes mellitus. Am J Obstet Gynecol 2010;203(6):556-e1-6.

184. Karcaaltincaba D, Yalvac S, Kandemir O, et al. Glycosylated hemoglobin level in the second trimester predicts birth weight and amniotic fluid volume in non-diabetic pregnancies with abnormal screening test. J Matern Fetal Neonatal Med 2010;23(10):1193-9.

185. Silva JC, Pacheco C, Bizato J, et al. Metformin compared with glyburide for the management of gestational diabetes. Int J Gynaecol Obstet 2010;111(1):37-40.

186. Anderberg E, Kallen K, Berntorp K. The impact of gestational diabetes mellitus on pregnancy outcome comparing different cut-off criteria for abnormal glucose tolerance. Acta Obstet Gynecol Scand 2010;89(12):1532-7.

187. Kvehaugen AS, Andersen LF, Staff AC. Anthropometry and cardiovascular risk factors in women and offspring after pregnancies complicated by preeclampsia or diabetes mellitus. Acta Obstet Gynecol Scand 2010;89(11):1478-85.

188. Pugh SK, Poole AT, Hill JB, et al. Abnormal 1 hour glucose challenge test followed by a normal 3 hour glucose tolerance test: does it identify adverse pregnancy outcome? J Miss State Med Assoc 2010;51(1):3-6.

189. Balaji V, Balaji MS, Alexander C, et al. Premixed insulin aspart 30 (Biasp 30) vs. premixed human insulin 30 (BHI 30) in gestational diabetes mellitus - a pilot study. J Assoc Physicians India 2010;58:99-101.

190. Barakat MN, Youssef RM, Al-Lawati JA. Pregnancy outcomes of diabetic women: charting Oman's progress towards the goals of the Saint Vincent Declaration. Ann Saudi Med 2010;30(4):265-70.

191. Negrato CA, Rafacho A, Negrato G, et al. Glargine vs. NPH insulin therapy in pregnancies complicated by diabetes: an observational cohort study. Diabetes Res Clin Pract 2010;89(1):46-51.

192. Fadl HE, Ostlund IK, Magnuson AF, et al. Maternal and neonatal outcomes and time trends of gestational diabetes mellitus in Sweden from 1991 to 2003. Diabetic Med 2010;27(4):436-41.

193. Flores-Le Roux JA, Chillaron JJ, Goday A, et al. Peripartum metabolic control in gestational diabetes. Am J Obstet Gynecol 2010;202(6):568-e1-6.

194. Geifman-Holtzman O, Machtinger R, Spiliopoulos M, et al. The clinical utility of oral glucose tolerance test at term: can it predict fetal macrosomia? Arch Gynecol Obstet 2010;281(5):817-21.

195. Buscicchio G, Gentilucci L, Giannubilo SR, et al. Computerized analysis of fetal heart rate in pregnancies complicated by gestational diabetes mellitus. Gynecol Endocrinol 2010;26(4):270-4.

196. Hamilton JK, Odrobina E, Yin J, et al. Maternal insulin sensitivity during pregnancy predicts infant weight gain and adiposity at 1 year of age. Obesity 2010;18(2):340-6.

197. Rowan JA, Gao W, Hague WM, et al. Glycemia and its relationship to outcomes in the metformin in gestational diabetes trial. Diabetes Care 2010;33(1):9-16.

198. Moore LE, Clokey D, Rappaport VJ, et al. Metformin compared with glyburide in gestational diabetes: a randomized controlled trial. Obstet Gynecol 2010;115(1):55-9.

199. Chodick G, Elchalal U, Sella T, et al. The risk of overt diabetes mellitus among women with gestational diabetes: A population-based study. Diabetic Med 2010;27(7):779-85.

200. Elkind-Hirsch KE, Ogden BW, Darensbourg CJ, et al. Clinical assessment of insulin action during late pregnancy in women at risk for gestational diabetes: Association of maternal glycemia with perinatal outcome. Int J Diabetes Mellitus 2010;2(1):3-9.

201. Perez-Ferre N, Galindo M, Fernandez MD, et al. The outcomes of gestational diabetes mellitus after a telecare approach are not inferior to traditional outpatient clinic visits. Int J Endocrinol 2010;386941.

202. Hedderson MM, Gunderson EP, Ferrara A. Gestational weight gain and risk of gestational

diabetes mellitus. Obstet Gynecol 2010;115(3):597-604.

203. O'Sullivan EP, Avalos G, O'Reilly M, et al. Atlantic Diabetes in Pregnancy (DIP): the prevalence and outcomes of gestational diabetes mellitus using new diagnostic criteria. Diabetologia 2011;54(7):1670-5.

204. Ehrlich SF, Crites YM, Hedderson MM, et al. The risk of large for gestational age across increasing categories of pregnancy glycemia. Am J Obstet Gynecol 2011;204(3):240-e1-6.

205. Corrado F, D'Anna R, Di VG, et al. The effect of myoinositol supplementation on insulin resistance in patients with gestational diabetes. Diabetic Med 2011;28(8):972-5.

206. Yogev Y, Melamed N, Chen R, et al. Glyburide in gestational diabetes--prediction of treatment failure. J Matern Fetal Neonatal Med 2011;24(6):842-6.

207. Ouzounian JG, Rosenheck R, Lee RH, et al. One-hour post-glucola results and pre-pregnancy body mass index are associated with the need for insulin therapy in women with gestational diabetes. J Matern Fetal Neonatal Med 2011;24(5):718-22.

208. Korpi-Hyovalti EA, Laaksonen DE, Schwab US, et al. Feasibility of a lifestyle intervention in early pregnancy to prevent deterioration of glucose tolerance. BMC Public Health 2011;11:179.

209. Riskin-Mashiah S, Damti A, Younes G, et al. Normal fasting plasma glucose levels during pregnancy: a hospital-based study. J Perinat Med 2011;39(2):209-11.

210. Kosus A, Kosus N, Turhan NO. Assessment of cardiomyopathy in fetuses of women with false positive oral glucose loading test. Eur J Obstet Gynecol Reprod Biol 2011;154(1):37-9.

211. Grant SM, Wolever TM, O'Connor DL, et al. Effect of a low glycaemic index diet on blood glucose in women with gestational hyperglycaemia. Diabetes Res Clin Pract 2011;91(1):15-22.

212. Deierlein AL, Siega-Riz AM, Chantala K, et al. The association between maternal glucose concentration and child BMI at age 3 years. Diabetes Care 2011;34(2):480-4.

213. Nanda S, Savvidou M, Syngelaki A, et al. Prediction of gestational diabetes mellitus by maternal factors and biomarkers at 11 to 13 weeks. Prenat Diagn 2011;31(2):135-41.

214. Yee LM, Cheng YW, Liddell J, et al. 50-Gram glucose challenge test: Is it indicative of outcomes in women without gestational diabetes mellitus? J Matern Fetal Neonatal Med 2011;24(9):1102-6.

215. Kaymak O, Iskender CT, Ustunyurt E, et al. Retrospective evaluation of perinatal outcome in women with mild gestational hyperglycemia. J Obstet Gynaecol 2011;37(8):986-91.

216. Jovanovic L, Pettitt D. Frequent monitoring of a1c during pregnancy as a treatment tool to guide therapy. Diabetes Technol Ther 2011;34(1):53-4.

217. Durnwald CP, Mele L, Spong CY, et al. Glycemic characteristics and neonatal outcomes of women treated for mild gestational diabetes. Obstet Gynecol 2011;117(4):819-27.

218. Ijas H, Vaarasmaki M, Morin-Papunen L, et al. Metformin should be considered in the treatment of gestational diabetes: a prospective randomised study. BJOG 2011;118(7):880-5.

219. Balaji V, Balaji M, Anjalakshi C, et al. Diagnosis of gestational diabetes mellitus in Asian-Indian women. Indian J Endocrinol Metab 2011;15(3):187-90.

220. Anderberg E, Landin-Olsson M, Kalen J, et al. Prevalence of impaired glucose tolerance and diabetes after gestational diabetes mellitus comparing different cut-off criteria for abnormal glucose tolerance during pregnancy. Acta Obstet Gynecol Scand 2011;90(11):1252-8.

221. Louie JC, Markovic TP, Perera N, et al. Randomized Controlled Trial Investigating the Effects of a Low-Glycemic Index Diet on Pregnancy Outcomes in Gestational Diabetes Mellitus. Diabetes Care 2011;34(11):2341-6.

222. Saxena P, Tyagi S, Prakash A, et al. Pregnancy outcome of women with gestational diabetes in a tertiary level hospital of north India. Indian J Community Med 2011;36(2):120-3.

223. Gandhi P, Bustani R, Madhuvrata P, et al. Introduction of metformin for gestational diabetes mellitus in clinical practice: has it had an impact? Eur J Obstet Gynecol Reprod Biol 2012;160(2):147-50.

224. Bahado-Singh RO, Mele L, Landon MB, et al. Fetal male gender and the benefits of treatment of mild gestational diabetes mellitus. Am J Obstet Gynecol 2012;206(5):422-5.

225. Dennedy MC, Avalos G, O'Reilly MW, et al. ATLANTIC-DIP: raised maternal body mass index (BMI) adversely affects maternal and fetal outcomes in glucose-tolerant women according to

International Association of Diabetes and Pregnancy Study Groups (IADPSG) criteria. J Clin Endocrinol Metab 2012;97(4):E608-E612.

226. Gasim T. Gestational diabetes mellitus: maternal and perinatal outcomes in 220 saudi women. Oman Med J 2012;27(2):140-4.

227. O'Dwyer V, Farah N, Hogan J, et al. Timing of screening for gestational diabetes mellitus in women with moderate and severe obesity. Acta Obstet Gynecol Scand 2012;91(4):447-51.

Excluded – Duplicate (N=10)

1. Cabero L, Corcoy R, Cerqueria MJ, Codina M. Treatment and Outcome of 100 Gestational Diabetics. 1989.

2. Retnakaran R. lsolated Hyperglycemia at 1 Hour on Oral Glucose Tolerance Test in Pregnancy Resembles Gestational Diabetes Mellitus in Predicting Postpartum Metabolic Dysfunction. Diabetes Care 2008;2008(7):1275-81.

3. Landon MB, Thom E, Spong CY, et al. The National Institute of Child Health and Human Development Maternal-Fetal Medicine Unit Network randomized clinical trial in progress: Standard therapy versus no therapy for mild gestational diabetes. Diabetes Care 2007;30(Suppl 2):S194-S199.

4. Landon MB. The national institute of child health and human development maternal-fetal medicine unit network randomized clinical trial in progress : Standard therapy versus no therapy for mild gestational diabetes. Proceedings of the fifth International Workshop-Conference on Gestational Diabetes Mellitus, 11-13 November 2005, Chicago, Illinois. Diabetes Care 2007;30(Suppl 2):S194-S199.

5. Landon MB, Thom E, Spong CY, Carpenter M, et al. The National Institute of Child Health and Human Development Maternal-Fetal Medicine Unit Network randomized clinical trial in progress: standard therapy versus no therapy for mild gestational diabetes. Chicago, IL 2007 p. S194-S199.

6. Langer O, Yogev Y, Most O. Gestational diabetes mellitus: the consequences of not treating. Am J Obstet Gynecol 2005;192(4):989-997.

7. Reichelt AJ, Spichler ER, Branchtein L, et al. Easting plasma glucose is a useful test for the detection of gestational diabetes. Diabetes Care 1998;21(8):1246-9.

8. Naylor CD, Sermer M, Chen EL, et al. Selective screening for gestational diabetes mellitus. N Engl J Med 1997;337(22):1591-6.

9. Balaji V, Balaji MS, Alexander C, et al. Premixed insulin aspart 30 (biasp 30) vs premixed human insulin 30 (bhi 30) in gestational diabetes mellitus - a pilot study. J Assoc Physicians India 2010;58(2):95-7.

10. Retnakaran R. -Cell Function Declines Within the First Year Postpartum in Women With Recent Glucose Intolerance in Pregnancy. Diabetes Care 2010;2010(8):1798-804.

Excluded – Intervention (N=12)

1. Persily CA. Relationships between the perceived impact of gestational diabetes mellitus and treatment adherence. J Obstet Gynecol Neonatal Nurs 1996;25(7):601-7.

2. Fisher JE, Smith RS, Lagrandeur R, et al. Gestational diabetes mellitus in women receiving beta-adrenergics and corticosteroids for threatened preterm delivery. Obstet Gynecol 1997;90(6):880-3.

3. Moses RG, Lucas EM, Knights S. Gestational diabetes mellitus. At what time should the postprandial glucose level be monitored? Aust N Z J Obstet Gynaecol 1999;39(4):457-60.

4. Nachum Z, Ben-Shlomo I, Weiner E, et al. Twice daily versus four times daily insulin dose regimens for diabetes in pregnancy: randomised controlled trial. BMJ: British Medical Journal (International Edition) 1999;319(7219):1223-7.

5. Holt RI, Goddard JR, Clarke P, et al. A postnatal fasting plasma glucose is useful in determining which women with gestational diabetes should

undergo a postnatal oral glucose tolerance test. Diabetic Med 2003;20(7):594-8.

6. Mosca A, Paleari R, Dalfra MG, et al. Reference intervals for hemoglobin A1c in pregnant women: data from an Italian multicenter study. Clin Chem 2006;52(6):1138-43.

7. Wong ML, Wong WH, Cheung YF. Fal myocardial performance in pregnancies complicated by gestational impaired glucose toleranceet. Ultrasound Obstet Gynecol 2007;29(4):395-400.

8. Kakad R, Anwar A, Dyer P, et al. Fasting plasma glucose is not sufficient to detect ongoing glucose intolerance after pregnancy complicated by gestational diabetes. Exp Clin Endocrinol Diabetes 2010;118(4):234-6.

9. Halkoaho A, Kavilo M, Pietila AM, et al. Does gestational diabetes affect women's health-related quality of life after delivery? Eur J Obstet Gynecol Reprod Biol 2010;148(1):40-3.

10. Karcaaltincaba D, Buyukkaragoz B, Kandemir O, et al. Gestational diabetes and gestational impaired glucose tolerance in 1653 teenage pregnancies: prevalence, risk factors and pregnancy outcomes. J Pediatr Adolesc Gynecol 2011;24(2):62-5.

11. Perera NJ, Molyneaux L, Constantino MI, et al. Suboptimal performance of blood glucose meters in an antenatal diabetes clinic. Diabetes Care 2011;34(2):335-7.

12. Catalano PM, McIntyre HD, Cruickshank JK, et al. The hyperglycemia and adverse pregnancy outcome study: associations of GDM and obesity with pregnancy outcomes. Diabetes Care 2012;35(4):780-6.

Excluded – Key Question 1 RCS (N=54)

1. Swinn RA, Wareham NJ, Gregory R, et al. Excessive secretion of insulin precursors characterizes and predicts gestational diabetes. Diabetes 1995;44(8):911-5.

2. Hooper DE. Detecting GD and preeclampsia. Effectiveness of routine urine screening for glucose and protein. J Reprod Med 1996;41(12):885-8.

3. Landy HJ, Gomez-Marin O, O'Sullivan MJ. Diagnosing gestational diabetes mellitus: use of a glucose screen without administering the glucose tolerance test. Obstet Gynecol 1996;87(3):395-400.

4. Tan YY, Yeo GS. Impaired glucose tolerance in pregnancy--is it of consequence? Aust N Z J Obstet Gynaecol 1996;36(3):248-55.

5. Kousta E, Lawrence NJ, Penny A, et al. Implications of new diagnostic criteria for abnormal glucose homeostasis in women with previous gestational diabetes. Diabetes Care 1999;22(6):933-7.

6. Atilano LC, Lee-Parritz A, Lieberman E, et al. Alternative methods of diagnosing gestational diabetes mellitus. Am J Obstet Gynecol 1999;181(5 Pt 1):1158-61.

7. Khine ML, Winklestein A, Copel JA. Selective screening for gestational diabetes mellitus in adolescent pregnancies. Obstet Gynecol 1999;93(5 Pt 1):738-42.

8. Jimenez-Moleon JJ, Bueno-Cavanillas A, Luna-del-Castillo JD, et al. Predictive value of a screen for gestational diabetes mellitus: influence of associated risk factors. Acta Obstet Gynecol Scand 2000;79(11):991-8.

9. Grover J, Beall MH, Ross MG. Intrapartum screen for diabetes in patients without prenatal care: use of labor admission serum glucose. J Matern Fetal Med 2000;9(4):216-8.

10. Shamsuddin K, Mahdy ZA, Siti R, I, et al. Risk factor screening for abnormal glucose tolerance in pregnancy. Int J Gynaecol Obstet 2001;75(1):27-32.

11. Kyle CV, Cundy TF. Screening for gestational diabetes mellitus: can we be more efficient? Aust N Z J Obstet Gynaecol 2001;41(3):285-90.

12. Chan LY, Wong SF, Ho LC. Diabetic family history is an isolated risk factor for gestational diabetes after 30 years of age. Acta Obstet Gynecol Scand 2002;81(2):115-7.

13. Larijani B, Hossein-nezhad A, Rizvi SW, et al. Cost analysis of different screening strategies for gestational diabetes mellitus. Endocrine Pract 2003;9(6):504-9.

14. Miyakoshi K, Tanaka M, Ueno K, et al. Cutoff value of 1 h, 50 g glucose challenge test for screening of gestational diabetes mellitus in a Japanese population. Diabetes Res Clin Pract 2003;60(1):63-7.

15. De Sereday MS, Damiano MM, Gonzalez CD, et al. Diagnostic criteria for gestational diabetes in relation to pregnancy outcome. J Diabetes Complications 2003;17(3):115-9.

16. Schytte T, Jorgensen LG, Brandslund I, et al. The clinical impact of screening for gestational diabetes. Clin Chem Lab Med 2004;42(9):1036-42.

17. Jakobi P, Solt I, Weissman A. A 2 hour versus the 3 hour 100 g glucose tolerance test for diagnosing gestational diabetes mellitus. J Perinat Med 2004;32(4):320-2.

18. Caliskan E, Kayikcioglu F, Ozturk N, et al. A population-based risk factor scoring will decrease unnecessary testing for the diagnosis of gestational diabetes mellitus. Acta Obstet Gynecol Scand 2004;83(6):524-30.

19. Sun JH, See LC, Chiu TH, et al. An appropriate indicator for diagnosing gestational diabetes. Chang Gung Medical Journal 2005;28(12):824-8.

20. Agarwal MM, Dhatt GS, Punnose J, et al. Gestational diabetes: dilemma caused by multiple international diagnostic criteria. Diabetic Med 2005;22(12):1731-6.

21. Esakoff TF, Cheng YW, Caughey AB. Screening for gestational diabetes: different cut-offs for different ethnicities? Am J Obstet Gynecol 2005;193(3 Suppl):1040-4.

22. Dabelea D, Snell-Bergeon JK, Hartsfield CL, et al. Increasing prevalence of gestational diabetes mellitus (GDM) over time and by birth cohort: Kaiser Permanente of Colorado GDM Screening Program. Diabetes Care 2005;28(3):579-84.

23. Cheng YW, Esakoff TF, Block-Kurbisch I, et al. Screening or diagnostic: markedly elevated glucose loading test and perinatal outcomes. J Matern Fetal Neonatal Med 2006;19(11):729-34.

24. Rodacki M, Lacativa PG, Lima GA, et al. Can we simplify the 100-g oral glucose tolerance test in pregnancy? Diabetes Res Clin Pract 2006;71(3):247-50.

25. Agarwal MM, Dhatt GS, Punnose J. Gestational diabetes: an alternative, patient-friendly approach for using the diagnostic 100-g OGTT in high-risk populations. Arch Gynecol Obstet 2006;273(6):325-30.

26. Fadl H, Ostlund I, Nilsson K, et al. Fasting capillary glucose as a screening test for gestational diabetes mellitus. BJOG 2006;113(9):1067-71.

27. Johnston-MacAnanny EB, Ness A, Weinstein L. Diagnosis of gestational diabetes mellitus: is it time for a new critical value? J Reprod Med 2007;52(6):463-6.

28. Hackmon R, James R, O'Reilly GC, et al. The impact of maternal age, body mass index and maternal weight gain on the glucose challenge test in pregnancy. J Matern Fetal Neonatal Med 2007;20(3):253-7.

29. Khan HA, Sobki SH, Alhomida AS, et al. Indian J Clin Biochem 2007;22(1):65-70.

30. Ogonowski J, Miazgowski T, Homa K, et al. Low predictive value of traditional risk factors in identifying women at risk for gestational diabetes. Acta Obstet Gynecol Scand 2007;86(10):1165-70.

31. Rudge MV, Lima CA, Paulette TA, et al. Influence of lower cutoff values for 100-g oral glucose tolerance test and glycemic profile for identification of pregnant women at excessive fetal growth risk. Endocrine Pract 2008;14(6):678-85.

32. Yamasmit W, Chaithongwongwatthana S, Uerpairojkit B. A 50-g glucose challenge test: is there any diagnostic cut-off? J Med Assoc Thai 2008;91(9):1309-12.

33. Punthumapol C, Tekasakul P. 50 grams glucose challenge test for screening of gestational diabetes mellitus in each trimester in potential diabetic pregnancy. J Med Assoc Thai 2008;91(6):787-93.

34. Montagnana M, Lippi G, Targher G, et al. Glucose challenge test does not predict gestational diabetes mellitus. Intern Med 2008;47(13):1171-4.

35. Korucuoglu U, Biri A, Turkyilmaz E, et al. Glycemic levels with glucose loading test during pregnancy and its association with maternal and perinatal outcomes. Diabetes Res Clin Pract 2008;80(1):69-74.

36. Boriboonhirunsarn D, Sunsaneevithayakul P. Abnormal results on a second testing and risk of gestational diabetes in women with normal baseline glucose levels. Int J Gynaecol Obstet 2008;100(2):147-53.

37. Phaloprakarn C, Tangjitgamol S. Diagnosis of gestational diabetes mellitus using a modified 100 g oral glucose tolerance test. J Perinatol 2008;28(1):7-11.

38. Aldasouqi SA, Solomon DJ, Bokhari SA, et al. Glycohemoglobin A1c: A promising screening tool in gestational diabetes mellitus. Int J Diabetes Dev Ctries 2008;28(4):121-4.

39. Phaloprakarn C, Tangjitgamol S, Manusirivithaya S. A risk score for selective screening for gestational diabetes mellitus. Eur J Obstet Gynecol Reprod Biol 2009;145(1):71-5.

40. Wong VW, Garden F, Jalaludin B. Hyperglycaemia following glucose challenge test during pregnancy: When can a screening test become diagnostic? Diabetes Res Clin Pract 2009;83(3):394-6.

41. Karcaaltincaba D, Kandemir O, Yalvac S, et al. Prevalence of gestational diabetes mellitus and gestational impaired glucose tolerance in pregnant women evaluated by National Diabetes Data Group and Carpenter and Coustan criteria. Int J Gynaecol Obstet 2009;106(3):246-9.

42. Flack JR, Ross GP, Ho S, et al. Recommended changes to diagnostic criteria for gestational diabetes: impact on workload. Aust N Z J Obstet Gynaecol 2010;50(5):439-43.

43. Ruangvutilert P, Chaemsaithong P, Ruangrongmorakot K, et al. Development of a modified 100-gram oral glucose tolerance test for diagnosis of gestational diabetes mellitus and its diagnostic accuracy. J Med Assoc Thai 2010;93(10):1121-7.

44. Kalamegham R, Nuwayhid BS, Mulla ZD. Prevalence of gestational fasting and postload single dysglycemia in Mexican-American women and their relative significance in identifying carbohydrate intolerance. Am J Perinatol 2010;27(9):697-704.

45. Agarwal MM, Dhatt GS, Shah SM. Gestational Diabetes Mellitus Simplifying the International Association of Diabetes and Pregnancy diagnostic algorithm using fasting plasma glucose. Diabetes Care 2010;33(9):2018-20.

46. Hansarikit J, Manotaya S. Sensitivity and specificity of modified 100-g oral glucose tolerance tests for diagnosis of gestational diabetes mellitus. J Med Assoc Thai 2011;94(5):540-4.

47. Teh WT, Teede HJ, Paul E, et al. Risk factors for gestational diabetes mellitus: implications for the application of screening guidelines. Aust N Z J Obstet Gynaecol 2011;51(1):26-30.

48. Shah A, Stotland NE, Cheng YW, et al. The association between body mass index and gestational diabetes mellitus varies by race/ethnicity. Am J Perinatol 2011;28(7):515-20.

49. Samuel A, Simhan HN. Clinical indications for abnormal early gestational 50-g glucose tolerance testing. Am J Perinatol 2011;28(6):485-7.

50. Huynh J, Ratnaike S, Bartalotta C, et al. Challenging the glucose challenge test. Aust N Z J Obstet Gynaecol 2011;51(1):22-5.

51. Church D, Halsall D, Meek C, et al. Random Blood Glucose Measurement at Antenatal Booking to Screen for Overt Diabetes in Pregnancy: A retrospective study. Diabetes Care 2011;34(10):2217-9.

52. Gandhi P, Farrell T. Gestational diabetes mellitus (GDM) screening in morbidly obese pregnant women. Eur J Obstet Gynecol Reprod Biol 2011;159(2):329-32.

53. Kosus A, Kosus N, Turhan NO. Gestational diabetes: comparision of the carpenter and the coustan thresholds with the new thresholds of turkish women and implications of variations in diagnostic criteria. J Matern Fetal Neonatal Med 2011.

54. Teede HJ, Harrison CL, Teh WT, et al. Gestational diabetes: Development of an early risk prediction tool to facilitate opportunities for prevention. Aust N Z J Obstet Gynaecol 2011;51(6):499-504.

Excluded – Outcome (N=34)

1. Tan YY, Liauw PC, Yeo GS. Using glucose tolerance test results to predict insulin requirement in women with gestational diabetes. Aust N Z J Obstet Gynaecol 1995;35(3):262-6.

2. Phillipov G. Short- and long-term reproducibility of the 1-h 50-g glucose challenge test. Clin Chem 1996;42(2):255-7.

3. Giampietro O, Matteucci E. Gestational diabetes mellitus (GDM) and macrosomia: a controversial story. Ann Ist Super Sanita 1997;33(3):399-402.

4. Kerbel D, Glazier R, Holzapfel S, et al. Adverse effects of screening for gestational diabetes: a prospective cohort study in Toronto, Canada. J Med Screen 1997;4(3):128-32.

5. Weiss PAM. Toward universal criteria for gestational diabetes : Relationships between

seventy-five and one hundred gram glucose loads and between capillary and venous glucose concentrations. Am J Obstet Gynecol 1998;1998(4):835.

6. Mills JL, Jovanovic L, Knopp R, et al. Physiological reduction in fasting plasma glucose concentration in the first trimester of normal pregnancy: the diabetes in early pregnancy study. Metabolism 1998;47(9):1140-4.

7. Stulberg RA, John SL, Houlden RL. Gestational age at screening, diagnosis and management of gestational diabetes in a Canadian community. Canadian Journal of Diabetes 1999;23(3):27-31.

8. Griffin ME, Coffey M, Johnson H, et al. Universal vs. risk factor-based screening for gestational diabetes mellitus: detection rates, gestation at diagnosis and outcome. Diabetic Med 2000;17(1):26-32.

9. Juutinen J, Hartikainen AL, Bloigu R, et al. A retrospective study on 435 women with gestational diabetes: Fasting plasma glucose is not sensitive enough for screening but predicts a need for insulin treatment [7]. Diabetes Care 2000;23(12):1858-9.

10. Weiss PA, Haeusler M, Tamussino K, et al. Can glucose tolerance test predict fetal hyperinsulinism? BJOG 2000;107(12):1480-5.

11. Zargar AH, Khan AK, Masoodi SR, et al. Prevalence of type 2 diabetes mellitus and impaired glucose tolerance in the Kashmir Valley of the Indian subcontinent. Diabetes Res Clin Pract 2000;47(2):135-46.

12. Kirwan JP, Huston-Presley L, Kalhan SC, et al. Clinically useful estimates of insulin sensitivity during pregnancy: validation studies in women with normal glucose tolerance and gestational diabetes mellitus. Diabetes Care 2001;24(9):1602-7.

13. Davey RX, Hamblin PS. Selective versus universal screening for gestational diabetes mellitus: an evaluation of predictive risk factors. Med J Aust 2001;174(3):118-21.

14. Jimenez-Moleon JJ, Bueno-Cavanillas A, Luna-del-Castillo JD, et al. Prevalence of gestational diabetes mellitus: variations related to screening strategy used. Eur J Endocrinol 2002;146(6):831-7.

15. Jorgensen LG, Schytte T, Brandslund I, et al. Fasting and post-glucose load--reference limits for peripheral venous plasma glucose concentration in pregnant women. Clin Chem Lab Med 2003;41(2):187-99.

16. Roggenbuck LF, Kleinwechter HJ, Demandt N, et al. Diagnostics of gestational diabetes: which cutoff-values are valid for capillary whole blood? Clin Lab 2004;50(7-8):403-8.

17. Bito T, Foldesi I, Nyari T, et al. Prediction of gestational diabetes mellitus in a high-risk group by insulin measurement in early pregnancy. Diabetic Med 2005;22(10):1434-9.

18. Bito T, Nyari T, Kovacs L, et al. 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. Eur J Obstet Gynecol Reprod Biol 2005;121(1):51-5.

19. Cypryk K, Pertynska-Marczewska M, Szymczak W, et al. Evaluation of metabolic control in women with gestational diabetes mellitus by the continuous glucose monitoring system: a pilot study. Endocrine Pract 2006;12(3):245-50.

20. Thomas B, Ghebremeskel K, Lowy C, et al. Nutrient intake of women with and without gestational diabetes with a specific focus on fatty acids. Nutrition 2006;22(3):230-6.

21. Seshiah V, Balaji V, Balaji MS, et al. Glycemic level at the first visit and prediction of GDM. J Assoc Physicians India 2007;55:630-2.

22. van LM, Opmeer BC, Zweers EJ, et al. Estimating the risk of gestational diabetes mellitus: a clinical prediction model based on patient characteristics and medical history. BJOG 2010;117(1):69-75.

23. Zisser HC, Biersmith MA, Jovanovic LB, et al. Fetal risk assessment in pregnancies complicated by diabetes mellitus. J Diabetes Sci Technol 2010;4(6):1368-73.

24. Luo ZC, Delvin E, Fraser WD, et al. Maternal glucose tolerance in pregnancy affects fetal insulin sensitivity. Diabetes Care 2010;33(9):2055-61.

25. Akinci B, Celtik A, Yener S, et al. Prediction of developing metabolic syndrome after gestational diabetes mellitus. Fertil Steril 2010;93(4):1248-54.

26. Radaelli T, Farrell KA, Huston-Presley L, et al. Estimates of insulin sensitivity using glucose and C-Peptide from the hyperglycemia and adverse pregnancy outcome glucose tolerance test. Diabetes Care 2010;33(3):490-4.

27. Meltzer SJ, Snyder J, Penrod JR, et al. Gestational diabetes mellitus screening and diagnosis: a prospective randomised controlled trial comparing

costs of one-step and two-step methods. BJOG 2010;117(4):407-15.

28. Das S, Behera MK, Misra S, et al. Beta-cell function and insulin resistance in pregnancy and their relation to fetal development. Metab Syndr Relat Disord 2010;8(1):25-32.

29. Lopez Caudana AE, Lopez RR, Gonzalez VC, et al. Prediction of alterations in glucose metabolism by glucose and insulin measurements in early pregnancy. Arch Med Res 2011;42(1):70-6.

30. Perovic M, Garalejic E, Gojnic M, et al. Sensitivity and specificity of ultrasonography as a screening tool for gestational diabetes mellitus. J Matern Fetal Neonatal Med 2011.

31. Wendland EM, Duncan BB, Mengue SS, et al. Lesser than diabetes hyperglycemia in pregnancy is related to perinatal mortality: a cohort study in Brazil. BMC Pregnancy Childbirth 2011;11:92.

32. Gibson KS, Waters TP, Catalano PM. Maternal weight gain in women who develop gestational diabetes mellitus. Obstet Gynecol 2012;119(3):560-5.

33. Verhaeghe J, Van HE, Benhalima K, et al. Glycated hemoglobin in pregnancies at increased risk for gestational diabetes mellitus. Eur J Obstet Gynecol Reprod Biol 2012;161(2):157-62.

34. Werner EF, Pettker CM, Zuckerwise L, et al. Screening for gestational diabetes mellitus: are the criteria proposed by the international association of the Diabetes and Pregnancy Study Groups cost-effective? Diabetes Care 2012;35(3):529-35.

Excluded – Population (N=15)

1. Gribble RK, Fee SC, Berg RL. The value of routine urine dipstick screening for protein at each prenatal visit. Am J Obstet Gynecol 1995;173(1):214-7.

2. Chang CJ, Wu JS, Lu FH, et al. Fasting plasma glucose in screening for diabetes in the Taiwanese population. Diabetes Care 1998;21(11):1856-60.

3. Bor MV, Bor P, Cevik C. Serum fructosamine and fructosamine-albumin ratio as screening tests for gestational diabetes mellitus. Arch Gynecol Obstet 1999;262(3-4):105-11.

4. Rich-Edwards JW, Colditz GA, Stampfer MJ, et al. Birthweight and the risk for type 2 diabetes mellitus in adult women. Ann Intern Med 1999;130(4 Pt 1):278-84.

5. Gray-Donald K, Robinson E, Collier A, et al. Intervening to reduce weight gain in pregnancy and gestational diabetes mellitus in Cree communities: an evaluation. CMAJ (Ottawa) 2000;163(10):1247-51.

6. Ko GT, Chan JC, Tsang LW, et al. Combined use of fasting plasma glucose and HbA1c predicts the progression to diabetes in Chinese subjects. Diabetes Care 2000;23(12):1770-3.

7. Agarwal MM, Punnose J, Dhatt GS. Gestational diabetes: Implications of variation in post-partum follow-up criteria. Eur J Obstet Gynecol Reprod Biol 2004;113(2):149-53.

8. Akhlaghi F, Hamedi AB. Comparison of maternal and fetal/neonatal complications in gestational and pre-gestational diabetes mellitus. Acta Medica Iranica 2005;43(4):263-7.

9. Sandbaek A, Lauritzen T, Borch-Johnsen K, et al. The comparison of venous plasma glucose and whole blood capillary glucose in diagnoses of Type 2 diabetes: a population-based screening study. Diabet Med 2005;22(9):1173-7.

10. Kraemer J, Klein J, Lubetsky A, et al. Perfusion studies of glyburide transfer across the human placenta: implications for fetal safety. Am J Obstet Gynecol 2006;195(1):270-4.

11. Franks PW, Looker HC, Kobes S, et al. Gestational glucose tolerance and risk of type 2 diabetes in young Pima Indian offspring. Diabetes 2006;55(2):460-5.

12. Ratner RE, Christophi CA, Metzger BE, et al. Prevention of diabetes in women with a history of gestational diabetes: effects of metformin and lifestyle interventions. J Clin Endocrinol Metab 2008;93(12):4774-9.

13. Phillips LS, Ziemer DC, Kolm P, et al. Glucose challenge test screening for prediabetes and undiagnosed diabetes. Diabetologia 2009;52(9):1798-807.

14. Opara PI, Jaja T, Onubogu UC. Morbidity and mortality amongst infants of diabetic mothers

admitted into a special care baby unit in Port Harcourt, Nigeria. Ital J Pediatr 2010;36(1):77.

15. Ohno MS, Sparks TN, Cheng YW, et al. Treating mild gestational diabetes mellitus: a cost-effectiveness analysis. Am J Obstet Gynecol 2011;205(3):282-e1-7.

Excluded – Publication Type (N=106)

1. Joshi R, Bharadwaj A. Gestational diabetes screening: can hemoglobin A~1~C measurement replace the glucose challenge test? Obstet Gynecol 2002:99(4 Suppl 1):S93.

2. Forest JC, Masse J, Garrido-Russo M. Glucose tolerance test during pregnancy: the significance of one abnormal value. Clin Biochem 1994;27(4):299-304.

3. JovanovicPeterson L, Bevier W, Peterson CM. A Cost-Effective Program to Normalize Birth-Weight (Bwt) by Screening for and Treatment of Glucose-Intolerance of Pregnancy (Igt). Diabetes 1995;44:A258.

4. Dolci M, Bianchini G, Andreani G, et al. Screening and treatment of gestational diabetes (GDM): The experience of seven years. Diabetologia 1996;39(Suppl 1):769.

5. Pavlic Renar I, Tomic M, Horvat B, Metelko Z. Screening and intervention in gestational diabetes mellitus. Fourth meeting for the implementation of the St. Vincent declaration, Lisbon 1997 p. 55.

6. Roncaglia N, Arreghini A, Bellini P, Bertalero C, et al. Gestational diabetes mellitus: is therapy always necessary? Rome, 1997 p. 68.

7. Harder T, Plagemann A, Kohlhoff R, Rohde W. Overweight and obesity in children of mothers with long-term insulin-dependent diabetes or gestational diabetes. 1997.

8. Cypryk K, Wilczynski J, Penza G, Krekora M. Early detection of gestational diabetes (GDM) improves pregnancy outcome. 1997.

9. Bancroft K, Tuffnell DJ, Mason GC, et al. A randomized controlled study of the management of impaired glucose intolerance in pregnancy. Br J Obstet Gynaecol 1998;105(Suppl 17):53-4.

10. Pavlic Renar I, Tomic M, Horvat B, Metelko Z. Screening and intervention in gestational diabetes mellitus. 1998.

11. Persson B, Hanson U. Neonatal morbidities in gestational diabetes mellitus. Diabetes Care 1998;21 Suppl 2:B79-B84.

12. Jovanovic L. Optimization of insulin therapy in patients with gestational diabetes. Endocrine Pract 2000;6(1):98-100.

13. Hadden D. Evidence-based screening for gestational diabetes? Diabetic Med 2000;17(5):402-4.

14. Kitzmiller JL. Cost analysis of diagnosis and treatment of gestational diabetes mellitus. Clin Obstet Gynecol 2000;43(1):140-53.

15. Souvatzoglou ES, Anastasiou E, Alevizaki M, et al. Is there any cutoff point of HbA1c levels indicative of the need for insulin treatment in women with gestational diabetes? Diabetologia 2001;44(Suppl 1):A42.

16. de Aguiar LG, de Matos HJ, de Brito GM. Could fasting plasma glucose be used for screening high-risk outpatients for gestational diabetes mellitus? Diabetes Care 2001;24(5):954-5.

17. Berger H, Crane J, Farine D, et al. Screening for gestational diabetes mellitus. J Obstet Gynaecol Can 2002;24(11):894-912.

18. McElduff A, Hitchman R. Screening for gestational diabetes: The time of day is important. The Medical Journal Of Australia 2002;176(3):136.

19. Bustani RJ, Todd DM, Akinsola M, et al. Increased insulin usage and reduced macrosomia in gestational diabetes mellitus managed with post-prandial blood glucose targets. Diabetologia 2002;45(Suppl 2):A291.

20. Jang HC, Park B, Park J, et al. Carbohydrate restricted diet in Korean women with mild gestational diabetes mellitus. Diabetes 2002;51:A615.

21. Castracane VD, Myles TD, Driggs SC, White D, et al. Early Detection of Gestational Diabetes is Enhances With Glucose Tolerance Testing in Early Pregnancy Source. Los Angeles, CA 2002 p. 171A.

22. Seshiah V, Balaji V, Balaji MS. Diagnosis and management of diabetes in pregnancy. J Indian Med Assoc 2003;101(12):742.

23. Giuffrida FM, Castro AA, Atallah AN, et al. Diet plus insulin compared to diet alone in the treatment of gestational diabetes mellitus: a systematic review. Braz J Med Biol Res 2003;36(10):1297-300.

24. McElduff A. Shared care: gestational diabetes. Aust Fam Physician 2003;32(3):113-7.

25. Platt J, O'Brien W. Acarbose therapy for gestational diabetes: A retrospective cohort study. Am J Obstet Gynecol 2003;189(6):S107.

26. Yogev Y, Langer O, Rosenn B, et al. Glucose challenge test as a predictor for gestational diabetes in mexican american women. Am J Obstet Gynecol 2003;189(6 Suppl):S86.

27. Senanayake H, Ariyaratne H, Wijeratne S. Is there a place for a single value oral glucose tolerance test for the diagnosis of gestational diabetes mellitus? Ceylon Med J 2004;49(4):136.

28. Fotinos C, Dodson S. Does tight control of blood glucose in pregnant women with diabetes improve neonatal outcomes? J Fam Pract 2004;53(10):838-40.

29. Arnqvist HJ, Hanson U, Nystrom L, et al. A population based study (G-DISS) of diagnosis of gestational diabetes and pregnancy outcome. Importance of fasting blood glucose. Diabetologia 2004;47(Suppl 1):A351.

30. Cheung NW, Oats JJ, McIntyre HD. Australian carbohydrate intolerance study in pregnant women: implications for the management of gestational diabetes. Aust N Z J Obstet Gynaecol 2005;45(6):484-5.

31. Yang HX, Gao XL, Dong Y, et al. Analysis of oral glucose tolerance test in pregnant women with abnormal glucose metabolism. Chin Med J 2005;118(12):995-9.

32. Hawkins JS, Lo J, Casey B, et al. Pregnancy outcomes associated with early diagnosis of diet-treated gestational diabetes. Am J Obstet Gynecol 2005;193(6):S91.

33. Moore L, Clokey D, Robinson A. A randomized trial of metformin compared to glyburide in the treatment of gestational diabetes. Am J Obstet Gynecol 2005;193(6):S92.

34. Ramos G, Jacobson G, Kirby R, et al. Comparison of glyburide and insulin for the management of gestational diabetics with greatly elevated oral glucose challenge test and fasting hyperglycemia. Am J Obstet Gynecol 2005;193(6):S93.

35. Barbour LA, Kahn BF, Davies JK, et al. Effectiveness of glyburide as an alternative to treat gestational diabetes. Diabetes 2005;54:A672.

36. Ross G. Gestational diabetes. Aust Fam Physician 2006;35(6):392-6.

37. Seely EW. Does treatment of gestational diabetes mellitus affect pregnancy outcome? Nat Clin Pract Endocrinol Metab 2006;2(2):72-3.

38. Loomis L, Lee J, Tweed E. What is appropriate fetal surveillance for women with diet-controlled gestational diabetes? J Fam Pract 2006;55(3):238-40.

39. Chollet MB, Pettitt DJ. Treatment of gestational diabetes mellitus. Clin Diabetes 2006;24(1):35-6.

40. Simmons D, Wolmarans L, Cutchie W, et al. Gestational diabetes mellitus: Time for consensus on screening and diagnosis. N Z Med J 2006;119(1228).

41. Cortez J, Tarsa M, Agent S, et al. Randomized controlled trial of acarbose vs. placebo in the treatment of gestational diabetes [abstract]. Am J Obstet Gynecol 2006;195(6 Suppl 1):S149.

42. Kwik M, Seeho S, Morris J, et al. ACHOIS confirmed: Adverse perinatal outcomes in pregnancies complicated by mild untreated gestational diabetes. Diabetes 2006;55:A418.

43. Parikh RM, Joshi SR, Menon PS, et al. Intensive glycemic control in diabetic pregnancy with intrauterine growth restriction is detrimental to fetus. Med Hypotheses 2007;69(1):203-5.

44. Moss JR, Crowther CA, Hiller JE, et al. Costs and consequences of treatment of gestational diabetes mellitus - evaluation from the ACHOIS randomised trial. Journal of Paediatr Child Health 2007;43(Suppl 1):A28-A29.

45. Williams M, Nguyen H, Towner D. Use of the maternal serum screen to predict adverse maternal outcomes among pregnant diabetics. Am J Obstet Gynecol 2007;197(6):S158.

46. De Mendonca NIM, Brace-Well-Milnes TJJ, Kaushal R, et al. Gestational diabetes in a multiethnic London population; the demographics, treatment requirements and pregnancy outcomes

and the implications of the ACHOIS trial for them. Diabetologia 2007;50(Suppl 1):S387.

47. Lundberg GD. Metformin Trumps insulin in the treatment of gestational diabetes. Medscape J Med 2008;10(7):179.

48. Rowan JA, Hague WM, Gao W, et al. Metformin versus insulin for the treatment of gestational diabetes. Obstet Gynecol Surv 2008;63(10):616-8.

49. Tieu J, Crowther CA, Middleton P, et al. Screening for gestational diabetes mellitus for improving maternal and infant health. Cochrane Database of Systematic Reviews 2008;(3):CD007222.

50. Simmons D. Diagnosis of gestational diabetes mellitus - A comparison of two screening tests. Which is the way ahead? Nat Clin Pract Endocrinol Metab 2008;4(2):72-3.

51. Landon MB. A Prospective Multicenter Randomized Treatment Trial of Mild Gestational Diabetes (Gdm). Am J Obstet Gynecol 2008;199(6):S2.

52. Lee S, Pettker C, Funai E, et al. Is Lowering the Diagnostic Threshold for Gestational Diabetes (Gdm) Cost-Effective? Implications from the Hyperglycemia and Adverse Pregnancy Outcomes (Hapo) Trial. Am J Obstet Gynecol 2008;199(6):S199.

53. Riskin-Mashiah S, Auslander R. First Trimester Fasting Hyperglycemia and Adverse Pregnancy Outcomes. Am J Obstet Gynecol 2008;199(6):S206.

54. Khurana R, Kozak SE, Thompson DM. Which values of the 3 hour 100g oral glucose tolerance test (OGTT) predict who will require insulin for treatment of gestational diabetes (GDM)? Diabetes 2008;57:A755.

55. Cheng YW, Block-Kurbisch I, Lydell J, et al. A Different Diagnostic Strategy Using the 100gram, 3-Hour Glucose Tolerance Test for the Diagnosis of Gestational Diabetes Mellitus. Reprod Sci 2008;15(1 Suppl):245A.

56. Wilson N, Goodwin W, Thomas S, et al. The multidisciplinary diabetes-endocrinology clinic and postprandial blood glucose monitoring in the management of gestational diabetes: Impact on maternal and neonatal outcomes. Diabetic Med 2009;26:181.

57. Nayyar V, Tier A, Hooker J, et al. Pregnancy outcomes inpatients with gestational diabetes. Diabetic Med 2009;26:179.

58. Mansell A, Gouveia C, Braggins F, et al. Early screening for gestational diabetes is essential to detect undiagnosed impaired glucose tolerance and Type 2 diabetes in a high risk, ethnically-diverse population. Diabetic Med 2009;26:117-8.

59. Fontaine P, Schaller S, Lenne X, et al. Increasing incidence of abnormal glucose tolerance in women with gestational diabetes (GDM) or mild gestational diabetes (MGH) in France: DIAGEST 2 study. Diabetologia 2009;52(Suppl 1):S457.

60. Vambergue A, Schaller S, Lenne X, et al. Anthropometric characteristics at 11 years in children exposed to maternal gestational diabetes mellitus or mild gestational hyperglycaemia in France: DIAGEST 2 study. Diabetologia 2009;52(Suppl 1):S64.

61. Yakubovich N, Qi Y, Sermer M, et al. Screening glucose challenge test in pregnancy: Impact of family history of diabetes on the likelihood of a false-negative result. Canadian Journal of Diabetes 2009;33(3):221.

62. Strevens H, Ursing D, Landin-Olsson M. Safe and efficient reporting of blood glucose values during pregnancy. Int J Gynecol Obstet 2009;107(Suppl 2):S644.

63. Divakar H, Kumar N, Manyonda I. Diagnostic criteria influence prevalence rates for gestational diabetes: Implications for interventions in an Indian pregnant population. Int J Gynecol Obstet 2009;107(Suppl 2):S439.

64. Phylos B, Lindow S, Coetzee E. Reproducibility of 75 g oral glucose tolerance test in pregnancy in South African population. Int J Gynecol Obstet 2009;107(Suppl 2):S436-S437.

65. Ng JM, Masson EA, Allan BJ, et al. Post-natal follow up of patients with gestational diabetes: One year onward. Pract Diabetes Int 2009;26(3):98.

66. Durnwald C. Glycemic characteristics of women treated for mild gestational diabetes and perinatal outcomes. Am J Obstet Gynecol 2009;201(6):S107.

67. Al-Haddabi R, Scott H, O'connell C, et al. Screening for gestational diabetes: does a false positive glucose challenge test predict adverse pregnancy outcome? Am J Obstet Gynecol 2009;201(6):S110.

68. Yogev Y, Chen R, Hod M, et al. Associations with preeclampsia: lessons from the hyperglycemia and adverse pregnancy outcome (HAPO) study. Am J Obstet Gynecol 2009;201(6):S27.

69. Gillman MW, Oakey H, Baghurst P, et al. Effect of Treatment of Gestational Diabetes on Obesity in the Next Generation. Obesity 2009;17:S315.

70. Negrato CA, Teixeira MF, Silva CA, et al. Use of Insulin Glargine vs NPH vs Diet in Pregnant Women with Gestational Diabetes. Diabetes 2009;58:A641.

71. Kohzuma T, Koga M. Lucica GA-L glycated albumin assay kit: a new diagnostic test for diabetes mellitus. Mol Diagn Ther 2010;14(1):49-51.

72. Zera CA, Seely EW. Diabetes: Treatment of gestational diabetes reduces obstetric morbidity. Nat Rev Endocrinol 2010;6(2):69-70.

73. Rowan JA. Gestational diabetes-complications, management, outcomes. Reprod Fertil Dev 2010;22:7.

74. Mahdavian M, Hivert MF, Baillargeon JP, et al. Gestational diabetes mellitus: Simplifying the international association of diabetes and pregnancy diagnostic algorithm using fasting plasma glucose. Diabetes Care 2010;33(11):e145.

75. Driul L, Londero A, Citossi A, et al. Neonatal and maternal outcomes by gestational diabetes mellitus and impaired glucose tolerance: A retrospective analysis of our 6-years experience. Arch Gynecol Obstet 2010;282(Suppl 1):S73.

76. Azriel S, Garcia BA, Camao I, et al. Relationship between perinatal outcomes and thyroid-peroxidase antibodies (TPO) in a cohort of pregnant women with gestational diabetes (GD). Diabetologia 2010;53:S438.

77. O'Reilly MW, Avalos G, Dennedy MC, et al. ATLANTIC DIP: Persistent postpartum glucose intolerance in women with previous gestational diabetes along the Irish Atlantic seaboard. Diabetologia 2010;53:S430.

78. Anderberg E, Landin-Olsson M, Kalen J, et al. Prevalence of diabetes mellitus after pregnancy with gestational diabetes mellitus using different cut-off criteria for abnormal glucose tolerance. Diabetologia 2010;53:S153.

79. Onofriescu M, Nemescu D, Tirnoveanu M, et al. Obstetrical and neonatal outcomes of gestational diabetes mellitus. J Matern Fetal Neonatal Med 2010;23(S1):562.

80. Marco P, Pastor M, Snchez EC, et al. Positive predictive value of O'Sullivan test in pregnants women and incidence of gestational diabetes in torrecardenas hospital. J Matern Fetal Neonatal Med 2010;23(S1):469.

81. Martinez PS, Abdulhaj MM, Andres NP, et al. A randomized study comparing metformin and insulin in the treatment of gestational diabetes mellitus. interim results. J Matern Fetal Neonatal Med 2010;23(S1):381.

82. Gregorini ME, Pagani G, Moretti P, et al. Treatment and gestational outcome in patients with gestational diabetes mellitus (GDM) in relation to the way of diagnosis. J Matern Fetal Neonatal Med 2010;23(S1):330.

83. Jenum AK, Sletner L, Voldner N, et al. The STORK Groruddalen research programme: a population-based cohort study of gestational diabetes, physical activity, and obesity in pregnancy in a multiethnic population. Rationale, methods, study population, and participation rates. SCAND J PUBLIC HEALTH 2010;38(5 Suppl):60-70.

84. Comparative evaluation of fasting plasma glucose and one hour 50-g glucose challenge test in screening gestational diabetes mellitus. J Zanjan Univ Med Sci Health Serv 2010;18(71):1-9.

85. Gayle C, Germain S, Marsh MS, et al. Comparing pregnancy outcomes for intensive versus routine antenatal treatment of gestational diabetes based on a 75gram oral glucose tolerance test 2-hour blood glucose 7.8-8.9mmol/l. Diabetologia 2010;53(Suppl 1).

86. Napoli A, Festa C, Merola G, et al. Low glycaemic index and hypocaloric diet therapy versus conventional approach in gestational diabetes/one abnormal value in pregnancy, after medical nutritional therapy failure. Diabetologia 2010;53(Suppl 1).

87. Hadden DR, Metzger BE, Lowe LP, et al. Hyperglycaemia and Adverse Pregnancy Outcome (HAPO) Study: Frequency of gestational diabetes mellitus (GDM) at collaborating centers based on IADPSG consensus panel recommended criteria. Diabetologia 2010;53(Suppl 1):S9.

88. Metzger BE, Lowe LP, Dyer AR, et al. The Hyperglycemia & Adverse Pregnancy Outcome (HAPO) Study: Associations of Higher Levels of Maternal Glucose and BMI with Macrosomia: An Example of Diabesity. Diabetes 2010;59:A42.

89. Metzger BE, Lowe LP, Dyer AR, et al. The Hyperglycemia & Adverse Pregnancy Outcome (HAPO) Study: Perinatal Outcome in Pregnancies with GDM and Fasting Plasma Glucose (FPG) <= 4.4 mmol/l. Diabetes 2010;59:A43.

90. Maher N, Reidy F, Walsh J, et al. Gestational diabetes-early treatment without rescreening: does this affect the incidence of macrosomia? Ir J Med Sci 2010;179(Suppl 2):S76-S77.

91. Trivedi N, Wen E, Aguayo J, et al. Impact of Diagnostic Intervals in Gestational Diabetes on Glycemic Control and Pregnancy Outcomes. Reprod Sci 2010;17(3):208A.

92. Simmons D, McElduff A, McIntyre HD, et al. Gestational Diabetes Mellitus: NICE for the US? A comparison of the American Diabetes Association and the American College of Obstetricians and Gynecologists guidelines with the UK National Institute for Health and Clinical Excellence guidelines. Diabetes Care 2010;33(1):34-7.

93. Ayach W, Calderon IM, Rudge MV, et al. [Comparison between two gestational diabetes screening tests and the perinatal outcome] [Portuguese]. Rev Bras Ginecol Obstet 2010;32(5):222-8.

94. Trivedi N, Aguayo J, Agent S, et al. Gestational diabetes in multiple gestations: incidence and implications of early screening. Reprod Sci 2011;18(Suppl 3):144A.

95. Ma KK. The obstetrical and neonatal implications of a low value on the glucose screening test. Reprod Sci 2011;18(Suppl 3):142A.

96. Brass E, Sheeder J, Dugoff.L. Is there a benefit to screening adolescents for gestational diabetes. Reprod Sci 2011;18(Suppl 3):139A.

97. Bertini AM, Silva JC, Narciso DRR, et al. Comparative study between metformin and glibenclamide in the treatment of gestational diabetes mellitus. Diabetes Technol Ther 2011;13(2).

98. Berggren EK, Boggess KA, Funk MJ, et al. Perinatal outcomes associated with changing diagnostic criteria for gestational diabetes National Diabetes Data Group versus Carpenter-Coustan. Am J Obstet Gynecol 2011;204(1):S224.

99. Zollinger T, Contreras K, Kominiarek M. Large for gestational age infants and the 3-hour oral glucose tolerance test values in gestational diabetes: Is there a relationship? Am J Obstet Gynecol 2011;204(1 Suppl):S111.

100. de VM, Wang J, Ferguson C, et al. How does the degree of hyperglycemia recorded during glucose tolerance testing for gestational diabetes impact perinatal outcome? Am J Obstet Gynecol 2011;204(1 Suppl):S109.

101. Lee EJ, Kim YH, Kwon JY, et al. Obstetric outcomes with a false positive 1-hour glucose challenge test. Am J Obstet Gynecol 2011;204(1 Suppl):S107-S108.

102. Wen E, Trivedi N, Aguayo J, et al. Early versus routine diagnosis of gestational diabetes mellitus: Comparison of perinatal outcomes and postpartum screening. Am J Obstet Gynecol 2011;204(1 Suppl):S107.

103. Centre for Reviews and Dissemination. Screening for gestational diabetes mellitus (Structured abstract). Database of Abstracts of Reviews of Effects 2011;4.

104. Liu Y, Wang J, Du M. Analysis for gestational diabetes screening of 1 676 pregnant women. [Chinese]. Matern Child Health Care China 2011;26(19):-2921.

105. Gorriz S, Esteve S, Guerrero A. Usefulness of Screening for Gestational Diabetes. Clin Chem Lab Med 2011;49(Suppl 1):S379.

106. Klebanoff M. Treatment of Gestational Diabetes (Gdm), Weight Gain and Perinatal Outcome - Marginal Structural Model (Msm) Analysis. Am J Epidemiol 2011;173(11 Suppl):S41.

Excluded – Study Design (N=11)

1. Mires GJ, Williams FL, Harper V. Screening practices for gestational diabetes mellitus in UK obstetric units. Diabet Med 1999;16(2):138-41.

2. Kremer CJ, Duff P. Glyburide for the treatment of gestational diabetes. Am J Obstet Gynecol 2004;190(5):1438-9.

3. Tanir HM, Sener T, Gurer H, et al. A ten-year gestational diabetes mellitus cohort at a university clinic of the mid-Anatolian region of Turkey. Clin Exp Obstet Gynecol 2005;32(4):241-4.

4. Dunne F. Type 2 diabetes and pregnancy. [Review] [56 refs]. Semin Fetal Neonatal Med 2005;10(4):333-9.

5. Ferrara A, Weiss NS, Hedderson MM, et al. Pregnancy plasma glucose levels exceeding the American Diabetes Association thresholds, but below the National Diabetes Data Group thresholds

for gestational diabetes mellitus, are related to the risk of neonatal macrosomia, hypoglycaemia and hyperbilirubinaemia. Diabetologia 2007;50(2):298-306.

6. Gonzalez-Quintero VH, Istwan NB, Rhea DJ, et al. Antenatal factors predicting subsequent need for insulin treatment in women with gestational diabetes. J Womens Health 2008;17(7):1183-7.

7. Davenport MH, Mottola MF, McManus R, et al. A walking intervention improves capillary glucose control in women with gestational diabetes mellitus: a pilot study. Appl Physiol Nutr Metab 2008;33(3):511-7.

8. Son GH, Kwon JY, Kim YH, et al. Maternal serum triglycerides as predictive factors for large-for-gestational age newborns in women with gestational diabetes mellitus. Acta Obstet Gynecol Scand 2010;89(5):700-4.

9. Sapienza AD, Francisco RP, Trindade TC, et al. Factors predicting the need for insulin therapy in patients with gestational diabetes mellitus. Diabetes Res Clin Pract 2010;88(1):81-6.

10. Patel S, Fraser A, Davey SG, et al. Associations of gestational diabetes, existing diabetes, and glycosuria with offspring obesity and cardiometabolic outcomes. Diabetes Care 2012;35(1):63-71.

11. Renar IP, Tomic M, Horvat B, Metelko Z. Screening and intervention in getational diabetes mellitus. Diabetologia 1997;40(suppl 1):848.

Excluded – Unobtainable (N=7)

1. Bell AW. Insulin Resistance in Pregnancy: Implications for Gestational Diabetes Source: In: Nutrition Society of Australia; p. 11-19; The Society; 1995 Series: Proceedings – Nutrition Society of Australia. Number: Vol 19 ISSN: 0314-1004 Language: English. Melbourne, Australia.

2. Mori M, Dolci M, Baccetti F. Evaluation after 1, 2, 3 years to delivery of glucose tolerance in women with gestational diabetes and of sons' development anthropometric. 1997.

3. Meyer WJ, Carbone J, Gauthier DW, et al. Early gestational glucose screening and gestational diabetes. J Reprod Med 1996;41(9):675-9.

4. Gorgojo Martinez JJ, Almodovar Ruiz F, Lopez Hernandez E, et al. Incidence of gestational diabetes mellitus according to different diagnostic criteria in the southeast Madrid area. Influence of diagnosis on materno-fetal parameters. Rev Clin Esp 2002;202(3):136-41.

5. Fan ZT, Yang HX, Gao XL, et al. Pregnancy outcome in gestational diabetes. Int J Gynaecol Obstet 2006;94(1):12-6.

6. Eslamian L, Ramezani Z. Breakfast as a screening test for gestational diabetes. Int J Gynaecol Obstet 2007;96(1):34-5.

7. Cheng YW, Block-Kurbisch I, Lydell J, Caughey AB. 2008. (missing reference data)

8. Sultan M, Khlaif H. Impact of gestational impaired glucose tolerance test (GIGTT) on pregnancy outcome. Jamahiriya Med J 2010;10(4):268-71.

Appendix F. Key Question 1 – HSROC Curves

Hierarchical summary receiver-operator curves (HSROC) with the 95 percent confidence ellipse are shown below for two different comparisons. The summary graphic compares the sensitivity and specificity for all studies comparing a particular screening test with GDM diagnostic criteria. All points are clustered in the upper left hand quadrant and there is no overalp between the 95 percent confidence ellipse and the diagonal null line. This indicates that the ability of the screening test to correctly classify patients with GDM is significantly better than random classification.

Figure F-1. HSROC curve: 50 g OGCT (≥140 mg/dL and ≥130 mg/dL) by Carpenter-Coustan criteria

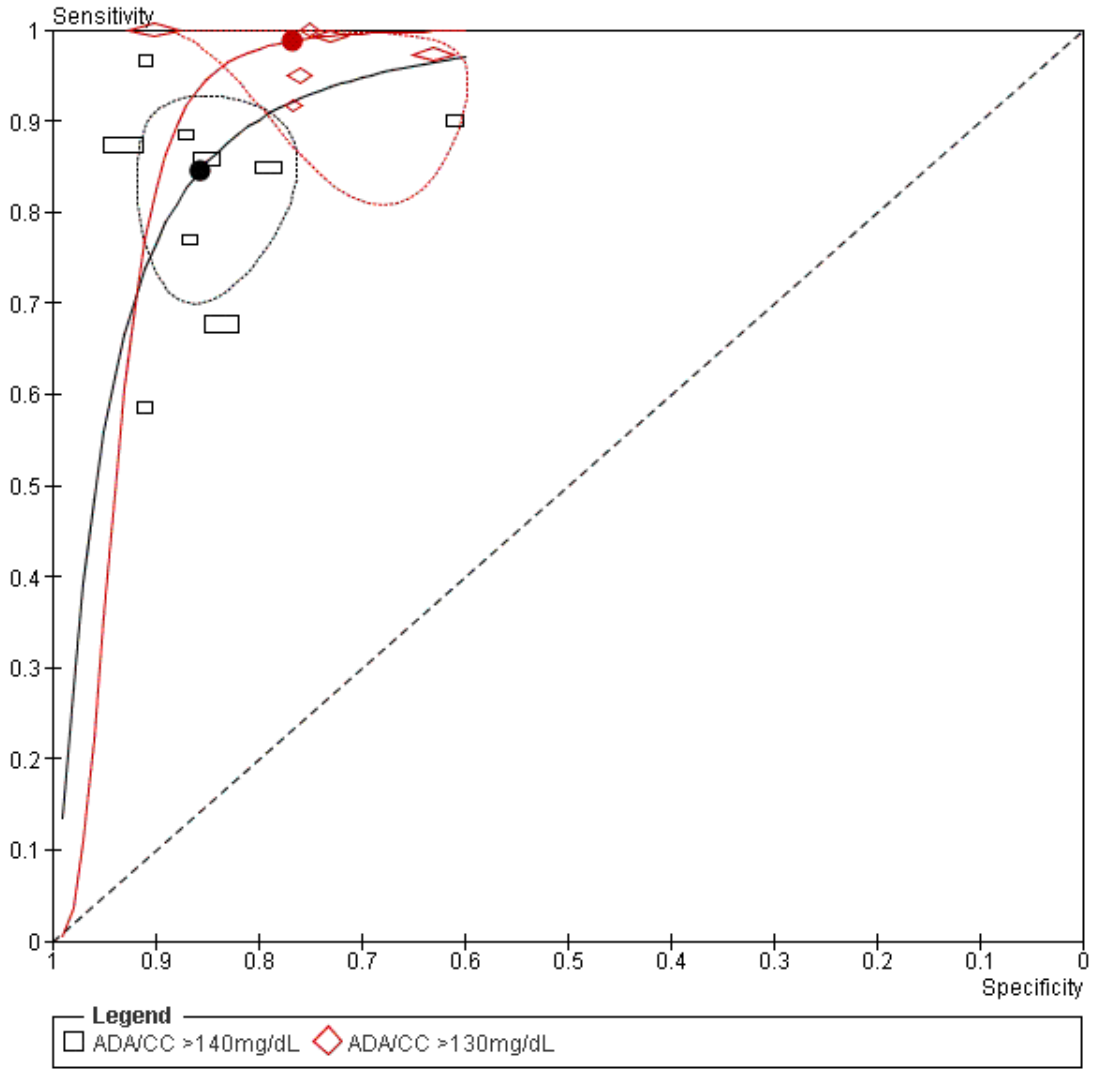
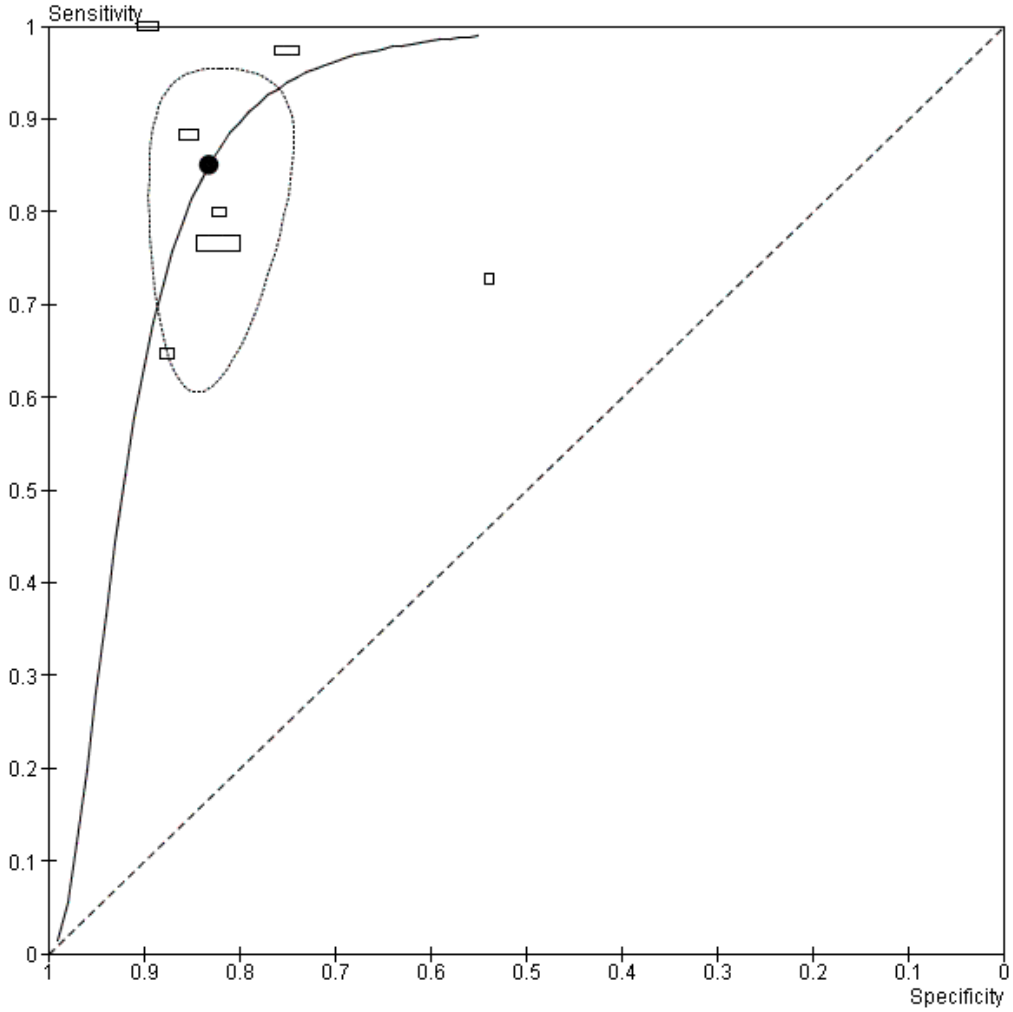


Figure F-2. HSROC curve: 50 g OGCT (≥ 140 mg/dL) by NDDG criteria



Appendix G. Adjusted Analyses for KQ3

Tables G-1 and G-2, on the following pages, provide unadjusted and adjusted results for maternal and offspring outcomes, respectively. The data that contributed to the meta-analysis for each comparison and outcome are provided. The data used in the meta-analyses and reported in the main report were unadjusted data from the relevant studies. We have also included the following for each study: whether the study provided adjusted results; what the adjusted effect estimate was (with its 95% confidence interval); whether the adjusted results were different from the unadjusted results in terms of statistical significance; and the variables that were controlled for in the adjusted analyses. For the overall pooled estimate within each comparison, we have noted whether the estimate would have changed if the adjusted values were used rather than the unadjusted values. For comparisons and outcomes with single studies, we have indicated whether the unadjusted and adjusted estimates differed in terms of statistical significance.

Table G-1. Maternal outcomes: Unadjusted data included in meta-analyses for Key Question 3 and adjusted effect estimates where available from included studies

Author, Year	n/N*	n/N*	Weight	Effect estimate (95% CI)†	Were there adjusted results?	Adjusted effect estimate (95% CI)	Adjusted results different	Variables in model	Impact of adjusted results on pooled estimates
PREECLAMPSIA									
CC GDM vs. no GDM									
Cheng, 2009	17/273	627/ 13,940	52.5%	1.38 (0.87, 2.21)	yes	1.3 (0.71, 2.38)	no	Parity, maternal age, race or ethnicity, gestational weight gain, gestational age at delivery, year of delivery, epidural anesthesia, induction of labor, (with mode of delivery and episiotomy additionally controlled for perineal laceration, postpartum hemorrhage, shoulder dystocia, and birth trauma)	
Naylor, 1996	10/115	144/2,940	30.4%	1.78 (0.96, 3.28)	no	n/a	n/a		
Pennison, 2001	9/43	10/69	17.2%	1.44 (0.64, 3.27)	yes	1.56 (0.58, 4.22)	no	African American race, elevated BMI	
Total (95% CI)	**431**	**16,949**	**100.0%**	**1.50 (1.07, 2.11)**					**Adding adjusted values would not change significance**
CC GDM vs. false-positive									
Berggren, 2011	58/460	264/3,117	86.8%	1.49 (1.14, 1.94)	yes	1.47 (1.02, 2.13)	no	Parity, maternal delivery age over 35 years, ethnicity, delivery year; cesarean and operative deliveries were also controlled for prior cesarean.	Summary measure is adjusted prevalence ratio
Naylor, 1996	10/115	31/580	13.2%	1.63 (0.82, 3.22)	no	n/a	n/a		
Total (95% CI)	**575**	**3,697**	**100.0%**	**1.51 (1.17, 1.93)**					**Adding adjusted values would not change significance**
NDDG 1 abnormal OGTT vs. no GDM									

Kim, 2002	5/122	18/577	100.0%	1.33 (0.48, 3.65)	no	n/a	n/a		
Total (95% CI)	**122**	**577**	**100.0%**	**1.33 (0.48, 3.65)**					**No change**
NDDG false-positive vs. no GDM									
Biri, 2009	7/326	21/1,432	35.5%	1.46 (0.63, 3.42)	no	n/a	n/a		
Stamilio, 2004	10/164	107/1,661	64.5%	0.95 (0.51, 1.77)	yes	0.33 (0.1, 1.11)	no	Body mass index, parity, gestational age at delivery, chronic hypertension, tobacco use, race, midtrimester serum -fetoprotein and human chorionic gonadotropin levels, maternal age, and history of preeclampsia in a prior pregnancy.	
Total (95% CI)	**490**	**3,093**	**100.0%**	**1.10 (0.67, 1.83)**					**Adding adjusted values would not change significance**
WHO IGT vs. no GDM									
Jensen, 2003	16/289	158/2,596	50.3%	0.91 (0.55, 1.50)	yes	0.9 (0.5,1.8)	no	Pre-pregnancy BMI, maternal age,parity, smoking, weight gain during pregnancy, gestational age, anamnestic risk indicators for GDM, ethnic background and clinical centre.	
Nord, 1995	13/223	14/391	42.1%	1.63 (0.78, 3.40)	no	n/a	n/a		
Yang, 2002	3/102	0/302	7.6%	20.59 (1.07, 395.30)	yes	2.1 (0.89, 4.94)	yes		
Total (95% CI)	**614**	**3,289**	**100.0%**	**1.47 (0.62, 3.52)**					**Adding adjusted values would not change significance**
MATERNAL HYPERTENSION									
CC vs. no GDM									
Chou, 2010	10/489	238/ 10,116	22.6%	0.87 (0.46, 1.63)	no	n/a	n/a		
Landon, 2011	62/455	31/423	34.1%	1.86 (1.23, 2.80)	yes	1.94 (1.09, 3.52)	no	Maternal age, gestational age at enrollment and at	

								delivery, parity, BMI, and race and ethnicity	
Lapolla, 2011	9/112	76/1,815	21.1%	1.92 (0.99, 3.73)	no	n/a	n/a		
Ricart, 2005	10/263	108/6,350	22.2%	2.24 (1.18, 4.22)	yes	2.34 (1.15, 4.77)	no	Maternal BMI, fetal sex (male), gestational age, maternal age, macrosomia (yes), PIH (yes)	
Total (95% CI)	**1,319**	**18,704**	**100.0%**	**1.64 (1.11, 2.42)**					**Adding adjusted values would not change significance**
CC vs. false-positive									
Berggren, 2011	33/460	150/3,117	77.6%	1.49 (1.04, 2.15)	yes	1.48 (1.02,2.13)	no	Parity, maternal delivery age over 35 years, ethnicity, delivery year; cesarean and operative deliveries were also controlled for prior cesarean.	Summary measure is adjusted prevalence ratio
Ricart, 2005	10/263	42/1,838	22.4%	1.66 (0.85, 3.28)	no	n/a	n/a		
Total (95% CI)	**723**	**4,955**	**100.0%**	**1.53 (1.11, 2.11)**					**Adding adjusted values would not change significance**
CC False-positive vs. no GDM									
Ricart, 2005	42/ 1,838	108/6,350	100.0%	1.35 (0.94, 1.94)	yes	1.25 (0.83, 1.90)	no	Maternal BMI, fetal sex (male), gestational age, maternal age, macrosomia (yes), PIH (yes)	**Adding adjusted values would not change significance**
CC 1 abnormal vs. no GDM									
Corrado, 2009	21/152	27/624	76.9%	3.19 (1.86, 5.49)	yes	2.3 (1.23,4.6)	no	Age and BMI (adjusted estimate for “hypertensive disorders)	
Vambergue, 2000	14/131	5/108	23.1%	2.31 (0.86, 6.21)	no	n/a	n/a		
Total (95% CI)	**283**	**732**	**100.0%**	**2.96 (1.84, 4.77)**					**Adding adjusted values would not change significance**
IADPSG GDM vs. no GDM									

Lapolla, 2011	9/112	76/1815	100.0%	1.92 (0.99, 3.73)	no	n/a	n/a		**No change**
IADPSG IGT (1 abnormal OGTT) vs no GDM									
Black, 2010	36/391	490/7,020	100.0%	1.32 (0.96, 1.82)	yes	1.49 (1.03, 2.16)	yes	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**Changed to statistically significant**
IADPSG IFG vs. no GDM									
Black, 2010	90/886	490/7,020	100.0%	1.46 (1.18, 1.80)	yes	1.29 (1.01, 1.66)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**
IADPST IGT-2 vs. no GDM									
Black, 2010	11/83	490/7,020	100.0%	1.90 (1.09, 3.31)	yes	2.33 (1.20, 4.51)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**
IADPSG IGT IFG vs no GDM									
Black, 2010	47/331	490/7,020	100.0%	2.03 (1.54, 2.69)	yes	2.01 (1.42, 2.84)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**
IADPSG IGT vs IFG									
Black, 2010	36/391	90/886	100.0%	0.91 (0.63, 1.31)	no	n/a	n/a		**No change**
IADPSG IGT vs. IGT-2									
Black, 2010	36/391	11/83	100.0%	0.69 (0.37, 1.31)	no	n/a	n/a		**No change**
IADPSG IGT vs IGT IFG									
Black, 2010	36/391	47/331	100.0%	0.65 (0.43, 0.98)	no	n/a	n/a		**No change**
IADPSG IFG vs. IGT-2									
Black, 2010	90/886	11/83	100.0%	0.77 (0.43, 1.37)	no	n/a	n/a		**No change**
IADPSG IFG vs IGT IFG									
Black, 2010	90/886	47/331	100.0%	0.72 (0.51, 0.99)	no	n/a	n/a		**No change**
IADPSG IGT-2 vs. IGT IFG									

Black, 2010	11/83	47/331	100.0%	0.93 (0.51, 1.72)	no	n/a	n/a		**No change**
WHO IGT vs. no GDM									
Jensen, 2003	16/289	158/2,596	158	0.91 (0.55, 1.50)	yes	0.9 (0.5,1.8)	no	pre-pregnancy BMI, maternal age, parity, smoking, weight gain during pregnancy, gestational age, anamnestic risk indicators for GDM, ethnic background and clinical centre.	**No change**
CESAREAN DELIVERY									
CC GDM vs. no GDM									
Cheng, 2009	62/ 273	2,356/ 13,940	13.2%	1.34 (1.08, 1.68)	yes	1.44 (1.01,2.07)	no	Parity, maternal age, race or ethnicity, gestational weight gain, gestational age at delivery, year of delivery, epidural anesthesia, induction of labor, (with mode of delivery and episiotomy additionally controlled for perineal laceration, postpartum hemorrhage, shoulder dystocia, and birth trauma)	
Chico, 2005	122/ 422	1,442/ 5,767	16.2%	1.16 (0.99, 1.35)	no	n/a	n/a		
Ching-Yu, 2010	196/ 489	3,761/ 10,116	18.2%	1.08 (0.96, 1.20)	no	n/a	n/a		
Langer, 2005	132/ 555	158/1,110	13.9%	1.67 (1.36, 2.06)	no	n/a	n/a		
Lapolla, 2011	49/112	564/1,815	13.3%	1.41 (1.13, 1.76)	no	n/a	n/a		
Naylor, 1996	34/115	585/2,940	10.5%	1.49 (1.11, 1.99)	yes	1.2 (0.7,2.0)	yes	Maternal age. race. parity. BMI, history of preeclampsia, history of cesarean delivery. gestational age. and current preeclampsia	
Pennison, 2001	13/43	17/69	3.9%	1.23 (0.66, 2.27)	yes	1.52 (0.54, 4.31)	no	African American race, elevated BMI	
Ricart, 2005	59/263	1,219/ 6,350	11.3%	1.17 (0.93, 1.47)	yes	0.95 (0.67, 1.35)	no	Maternal BMI, fetal sex (male), gestational age,	

								maternal age, macrosomia (yes), PIH (yes)	
Schwartz, 1999	38/154	1,110/ 7,207	10.8%	1.60 (1.21, 2.12)	no	n/a	n/a		
Total (95% CI)	**2,426**	**49,314**	**100.0%**	**1.32 (1.17, 1.48)**					**Adding adjusted results would likely reduce lower confidence interval closer to null; not sure whether significance would change**
CC GDM vs. false-positive									
Berggren, 2011	160/ 460	942/3,117	72.3%	1.15 (1.00, 1.32)	yes	1.16 (1.04, 1.30)	yes	Parity, maternal delivery age over 35 years, ethnicity, delivery year; cesarean and operative deliveries were also controlled for prior cesarean.	
Naylor, 1996	34/115	136/580	13.2%	1.26 (0.92, 1.73)	no	n/a	n/a		
Ricart, 2005	59/263	393/1,838	0.187%	1.05 (0.82, 1.34)	no	n/a	n/a		
Schwartz, 1999	38/154	197/1,066	14.5%	1.34 (0.99, 1.81)	no	n/a	n/a		
Total (95% CI)	**992**	**6,601**	**100.0%**	**1.16 (1.05, 1.29)**					**Adding adjusted values would not change significance**
CC GDM vs. 1 abnormal OGTT									
Chico, 2005	122/ 422	19/59	100.0%	0.90 (0.60, 1.34)	no	n/a	n/a		**No change**
CC 1 abnormal OGTT vs. no GDM									
Chico, 2005	19/59	1,442/ 5,767	15.5%	1.29 (0.89, 1.87)	no	n/a	n/a		
Corrado, 2009	85/152	243/624	73.1%	1.44 (1.21, 1.71)	yes	2.2 (1.55, 3.39)	no	Age and BMI	
Rust, 1996	14/78	32/205	6.6%	1.15 (0.65, 2.04)	no	n/a	n/a		
Vambergue, 2000	23/131	11/108	4.8%	1.72 (0.88, 3.37)	no	n/a	n/a		
Total (95% CI)	**420**	**6,704**	**100.0%**	**1.40 (1.21, 1.63)**					**Adding adjusted values would not**

									change significance
CC false-positive vs. no GDM									
Bo, 2004	103/ 315	28/91	4.0%	1.06 (0.75, 1.50)	no	n/a	n/a		
Lapolla, 2007	45/128	100/334	5.8%	1.17 (0.88, 1.56)	no	n/a	n/a		
Naylor, 1996	136/ 580	585/2,940	17.8%	1.18 (1.00, 1.39)	yes	1.2 (0.9, 1.5)	no	Maternal age. race. parity. BMI, history of preeclampsia, history of cesarean delivery. gestational age. and current preeclampsia	
Ricart, 2005	393/ 1,838	1,219/ 6,350	46.9%	1.11 (1.01, 1.23)	yes	1.06 (0.91, 1.23)	no	Maternal BMI, fetal sex (male), gestational age, maternal age, macrosomia (yes), PIH (yes)	
Schwartz, 1999	197/ 1,066	1,110/ 7,207	25.5%	1.20 (1.05, 1.38)	no	n/a	n/a		
Total (95% CI)	**3,927**	**16,922**	**100.0%**	**1.15 (1.07, 1.23)**					**Adding adjusted value may lower the lower confidence bound closer to null; not clear whether significance would change**
NDDG 1 Abnormal OGTT vs. no GDM									
Kim, 2002	27/122	83/577	100.0%	1.69 (1.04, 2.75)	no	n/a	n/a		**No change**
CC 1 abnormal OGTT vs. false-positive									
Kwik, 2007	46/156	61/197	50.7%	0.95 (0.69, 1.31)	no	n/a	n/a		
Lapolla, 2007	27/48	45/128	49.3%	1.60 (1.14, 2.25)	no	n/a	n/a		
Total (95% CI)	**204**	**325**	**100.0%**	**1.23 (0.73, 2.06)**					**No change**
NDDG GDM vs no GDM									
Adams, 1998	4/16	10/64	100.0%	1.60 (0.58, 4.45)	no	n/a	n/a		**No change**
NDDG false-positive vs. no GDM									
Ardawi, 2000	24/187	67/529	3.9%	1.01 (0.66, 1.57)	no	n/a	n/a		
Hillier, 2007	208/326	785/1,432	83.2%	1.16 (1.06, 1.28)	no	n/a	n/a		
Retnakaran, 2008	44/128	23/74	4.3%	1.11 (0.73, 1.68)	no	n/a	n/a		

Stamilio, 2004	39/164	286/1,661	8.6%	1.38 (1.03, 1.85)	yes	1.76 (0.99, 3.14)	yes	Body mass index, parity, gestational age at delivery, chronic hypertension, tobacco use, race, midtrimester serum -fetoprotein and human chorionic gonadotropin levels, maternal age, and history of preeclampsia in a prior pregnancy.	
Total (95% CI)	**805**	**3,696**	**100.0%**	**1.17 (1.08, 1.28)**					**Adding adjusted values would not change significance**
WHO IGT vs no GDM									
Aberg, 2001	12/131	249/4,526	7.0%	1.67 (0.96, 2.89)	no	n/a	n/a		
Jensen, 2003	54/289	450/2,596	26.4%	1.08 (0.84, 1.39)	yes	1 (0.7, 1.4)	no	Pre-pregnancy body mass index (BMI), maternal age,parity, smoking, weight gain during pregnancy, gestational age, anamnestic risk indicators for GDM, ethnic background and clinical centre.	
Nord, 1995	38/223	45/391	12.7%	1.48 (0.99, 2.21)	no	n/a	n/a		
Yang, 2002	75/102	199/302	53.9%	1.12 (0.97, 1.29)	no	n/a	n/a		
Total (95% CI)	**745**	**7,815**	**100.0%**	**1.18 (1.01, 1.37)**					**Adding adjusted values would not change significance**
IADPSG GDM vs no GDM									
Lapolla, 2011	9/112	76/1,815	100.0%	1.92 (0.99, 3.73)	no	n/a	n/a		**No change**
IADPSG IGT vs no GDM									
Black, 2010	69/391	1,112/ 7,020	100.0%	1.11 (0.89, 1.39)	yes	1.03 (0.77,1.38)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**

IADPSG IFG vs no GDM									
Black, 2010	179/ 886	1,112/ 7,020	100.0%	1.28 (1.11, 1.47)	yes	1.16 (0.95,1.41)	yes	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**Changed to not statistically significant**
IADPSG IGT-2 vs no GDM									
Black, 2010	19/83	1,112/ 7,020	100.0%	1.58 (0.94, 2.64)	yes	1.39 (0.78, 2.46)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**
IADPSG IGT IFG vs. no GDM									
Black, 2010	69/331	1,112/ 7,020	100.0%	1.32 (1.06, 1.63)	yes	1.36 (1.00,1.85)	yes	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**Changed to not statistically significant**
IADPSG IGT vs. IFG									
Black, 2010	69/391	179/886	100.0%	0.87 [0.68, 1.12]	no	n/a	n/a		**No change.**
IADPSG IGT vs. IGT-2									
Black, 2010	69/391	19/83	100.0%	0.77 (0.49, 1.21)	no	n/a	n/a		**No change.**
IADPSG IGT vs. IGT IFG									
Black, 2010	69/391	69/331	100.0%	0.85 (0.63, 1.14)	no	n/a	n/a		**No change.**
IADPSG IFG vs. IGT-2									
Black, 2010	179/886	19/83	100.0%	0.88 (0.58, 1.34)	no	n/a	n/a		**No change.**
IADPSG IFG vs. IGT IFG									
Black, 2010	179/886	69/331	100.0%	0.97 (0.76, 1.24)	no	n/a	n/a		**No change.**
IADPSG IGT-2 vs. IGT IFG									
Black, 2010	19/83	69/331	100.0%	1.10 (0.70, 1.72)	no	n/a	n/a		**No change.**
MATERNAL BIRTH TRAUMA									
CC GDM vs no GDM									
Cheng, 2009	31/273	1,255/ 13,940	100.0%	1.26 (0.90, 1.76)	yes	1.16 (0.73,1.86)	no	Parity, maternal age, race or ethnicity, gestational weight gain, gestational age at delivery, year of	**No change**

								delivery, epidural anesthesia, induction of labor, (with mode of delivery and episiotomy additionally controlled for perineal laceration, postpartum hemorrhage, shoulder dystocia, and birth trauma)	
CC GDM vs false-positive									
Berggren, 2011	14/460	118/3,117	100.0%	0.80 (0.47, 1.39)	yes	0.83 (0.48, 1.44)	no	Parity, maternal delivery age over 35 years, ethnicity, delivery year; cesarean and operative deliveries were also controlled for prior cesarean.	
NDDG GDM vs no GDM									
Adams, 1998	2/16	4/64	100.0%	2.00 (0.40, 9.97)	no	n/a	n/a		**No change**
MATERNAL WEIGHT GAIN									
CC 1 abnormal OGTT vs no GDM									
Rust, 1996	36/78	38/205	100.0%	2.49 (1.71, 3.62)	no	n/a	n/a		**No change**
WHO IGT vs. No GDM (data presented are mean (SD), n for each group; weight; and mean difference with 95% CI)									
Yang, 2002	15.4 (6.5), 102	15.4 (5.6), 302	100.0%	0.00 (-1.41, 1.41)	no	n/a	n/a		**No change**
IADPSG IGT vs. NO GDM (data presented are mean (SD), n for each group; weight; and mean difference with 95% CI)									
Black, 2010	27.1 (14.5), 391	29.0 (13.7), 7,020	100.0%	-1.90 (-3.37, -0.43)	no	n/a	n/a		**No change**
IADPSG IFG vs NO GDM (data presented are mean (SD), n for each group; weight; and mean difference with 95% CI)									
Black, 2010	27.8 (15.2), 886	29.0 (13.7), 7,020	100.0%	-1.20 (-2.25, -0.15)	no	n/a	n/a		**No change**
IADPSG IGT-2 vs. NO GDM (data presented are mean (SD), n for each group; weight; and mean difference with 95% CI)									
Black, 2010	26.4 (11.6), 83	29.0 (13.7), 7,020	100.0%	-2.60 (-5.12, -0.08)	no	n/a	n/a		**No change**
IADPSG IGT IFG vs. NO GDM (data presented are mean (SD), n for each group; weight; and mean difference with 95% CI)									
Black, 2010	27.8 (14.8), 331	29.0 (13.7), 7,020	100.0%	-1.20 (-2.83, 0.43)	no	n/a	n/a		**No change**

IADPSG IGT vs. IFG (data presented are mean (SD), n for each group; weight; and mean difference with 95% CI)								
Black, 2010	27.1 (14.5), 391	27.8 (15.2), 886	100.0%	-0.70 (-2.45, 1.05)	no	n/a	n/a	**No change**
IADPSG IGT vs IGT-2 (data presented are mean (SD), n for each group; weight; and mean difference with 95% CI)								
Black, 2010	27.1 (14.5), 391	26.4 (11.6), 83	100.0%	0.70 (-2.18, 3.58)	no	n/a	n/a	**No change**
IADPSG IGT vs. IGT IFG (data presented are mean (SD), n for each group; weight; and mean difference with 95% CI)								
Black, 2010	27.1 (14.5), 391	27.8 (14.8), 331	100.0%	-0.70 (-2.85, 1.45)	no	n/a	n/a	**No change**
IADPSG IFT vs. IGT-2 (data presented are mean (SD), n for each group; weight; and mean difference with 95% CI)								
Black, 2010	27.8 (15.2), 886	26.4 (11.6), 83	100.0%	1.40 (-1.29, 4.09)	no	n/a	n/a	**No change**
IADPSG IFG vs. IGT IFG (data presented are mean (SD), n for each group; weight; and mean difference with 95% CI)								
Black, 2010	27.8 (15.2), 886	27.8 (14.8), 331	100.0%	0.00 (-1.88, 1.88)	no	n/a	n/a	**No change**
IADPSG IGT-2 vs. IGT IFG (data presented are mean (SD), n for each group; weight; and mean difference with 95% CI)								
Black, 2010	26.4 (11.6), 83	27.8 (14.8), 331	100.0%	-1.40 (-4.36, 1.56)	no	n/a	n/a	**No change**
MATERNAL MORBIDITY/MORTALITY								
CC GDM vs no GDM								
Lapolla, 2011	26/112	299/1,815	100.0%	1.53 (0.97, 2.42)	no	n/a	n/a	**No change**
CC 1 ABNORMAL OGTT vs NO GDM								
Rust, 1996	5/78	13/205	100.0%	1.01 (0.37, 2.74)	no	n/a	n/a	**No change**
IADPSG GDM vs no GDM								
Lapolla, 2011	26/112	294/1,815	100.0%	1.43 (1.01, 2.04)	no	n/a	n/a	**No change**

* The information presented in these columns is number of patients with the outcome / numbers of patients per group, except where otherwise indicated.
† The effect estimates are risk ratios with 95% confidence intervals, unless otherwise indicated.
BMI = body mass index; CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; IADPSG = International Association of Diabetes and Pregnancy Study Groups; IFG = impaired fasting glucose; IFT = impaired fasting tolerance; IGT = impaired glucose tolerance; IGT-2 = double impaired glucose tolerance; NDDG = National Diabetes Data Group; n = number of patients with the outcome; N = numbers of patients per group; n/a = not applicable; OGTT = oral glucose tolerance test; PIH = Pregnancy induced hypertension;SD = standard deviation; WHO = World Health Organization

Table G-2. Offspring outcomes: Unadjusted data included in meta-analyses for Key Question 3 and adjusted effect estimates where available from included studies

Author, Year	n/N*	n/N*	Weight	Effect estimate (95% CI)[†]	Were there adjusted results?	Adjusted effect estimate (95% CI)	Adjusted results different	Variables in model	Impact of adjusted results on pooled estimates
Macrosomia >4,500 g									
CC GDM vs. no GDM									
Cheng, 2009	11/273	223/13,940	50.7%	2.52 (1.39, 4.56)	Yes	4.47 (2.26, 8.86)	no	Parity, maternal age, race or ethnicity, gestational weight gain, gestational age at delivery, year of delivery, epidural anesthesia, induction of labor, (with mode of delivery and episiotomy additionally controlled for perineal laceration, postpartum hemorrhage, shoulder dystocia, and birth trauma)	
Naylor, 1996	7/115	56/2,940	30.6%	3.20 (1.49, 6.86)	no	n/a	n/a		
Schwartz, 1999	4/91	108/4,190	18.7%	1.71 (0.64, 4.53)	no	n/a	n/a		
Total (95% CI)	**479**	**21,070**	**100.0%**	**2.52 (1.65, 3.84)**					**No difference in significance if adjusted estimate was added; may increase estimate of RR**
CC vs. false-positive									
Naylor, 1996	7/115	12/580	52.2%	2.94 (1.18, 7.31)	no	n/a	n/a		
Schwartz, 1999	4/91	28/605	47.8%	0.95 (0.34, 2.64)	no	n/a	n/a		
Total (95% CI)	**206**	**1185**	**100.0%**	**1.71 (0.56, 5.24)**					**No change**
CC false positive vs. no GDM									
Naylor, 1996	12/580	56/2940	39.0%	1.09 (0.59, 2.01)	no	n/a	n/a		
Schwartz, 1999	28/605	108/4,190	61.0%	1.80 (1.20, 2.70)	no	n/a	n/a		
Total (95% CI)	**1,185**	**7,130**	**100.0%**	**1.48 (0.91, 2.39)**					**No change**

NDDG GDM vs no GDM									
Adams, 1998	3/16	0/64	100.0%	26.76 (1.45, 493.62)	no	n/a	n/a		**No change**
Macrosomia >4,000 g									
CC GDM vs. no GDM									
Berkus, 1995	13/72	76/573	7.4%	1.36 (0.80, 2.32)	no	n/a	n/a		
Chico, 2005	22/422	288/5,767	10.1%	1.04 (0.68, 1.59)	no	n/a	n/a		
Chou, 2010	22/489	236/1,0116	10.0%	1.93 (1.26, 2.96)	no	n/a	n/a		
Hillier, 2007	25/173	905/7,609	11.8%	1.21 (0.84, 1.75)	no	n/a	n/a		
Langer, 2005	93/555	87/1,110	15.5%	2.14 (1.63, 2.81)	no	n/a	n/a		
Lapolla, 2011	12/112	145/1,815	7.0%	1.34 (0.77, 2.34)	no	n/a	n/a		
Naylor, 1996	33/115	395/2,940	14.3%	2.14 (1.58, 2.89)	no				
Pennison, 2001	6/43	5/69	2.2%	1.93 (0.63, 5.93)	no				
Ricart, 2005	21/263	292/6,350	10.0%	1.74 (1.13, 2.66)	yes	1.45 (0.83, 2.52)	yes	Maternal BMI, fetal sex (male), gestational age, maternal age, macrosomia (yes), PIH (yes)	
Schwartz, 1999	22/91	692/4,190	11.7%	1.46 (1.01, 2.12)	no	n/a	n/a		
Total (95% CI)	**2335**	**40,539**	**100.0%**	**1.61 (1.35, 1.92)**					**Adding adjusted value would not change significance**
CC GDM vs. false-positive									
Berggren, 2011	78/460	411/3,117	30.1%	1.29 (1.03, 1.60)	yes	1.25 (1.01,1.56)	no	Parity, maternal delivery age over 35 years, ethnicity, delivery year; cesarean and operative deliveries were also controlled for prior cesarean	
Hillier, 2007	25/173	122/999	17.3%	1.18 (0.79, 1.76)	no	n/a	n/a		
Naylor, 1996	33/115	80/580	20.0%	2.08 (1.46, 2.96)	no				
Ricart, 2005	21/263	131/1,838	15.2%	1.12 (0.72, 1.74)	no	n/a	n/a		
Schwartz, 1999	22/91	119/605	17.4%	1.23 (0.83, 1.83)	no	n/a	n/a		
Total (95% CI)	**1,102**	**7,139**	**100.0%**	**1.36 (1.10, 1.68)**					**Adding adjusted value would not change significance**

CC GDM vs 1 abnormal OGTT									
Berkus, 1995	13/72	18/87	31.1%	0.87 (0.46, 1.66)	no	n/a	n/a		
Chico, 2005	22/422	3/59	9.3%	1.03 (0.32, 3.32)	no	n/a	n/a		
Hillier, 2007	25/173	40/288	59.7%	1.04 (0.66, 1.65)	no	n/a	n/a		
Total (95% CI)	**667**	**434**	**100.0%**	**0.98 (0.69, 1.41)**					**No change**
CC 1 abnormal OGTT vs. no GDM									
Berkus, 1995	18/87	76/573/	20.8%	1.56 (0.98, 2.48)	no	n/a	n/a		
Chico, 2005	3/59	288/5,767	4.4%	1.02 (0.34, 3.08)	no	n/a	n/a		
Corrado, 2009	19/152	39/624	17.2%	2.00 (1.19, 3.36)	yes	2 (1.13, 3.61)	no	Age and BMI	
Hillier, 2007	40/288	905/7,609	39.0%	1.17 (0.87, 1.57)	no	n/a	n/a		
Lapolla, 2007	3/48	8/334	3.3%	2.61 (0.72, 9.50)	no	n/a	n/a		
Rust, 1996	6/78	18/205	6.7%	0.88 (0.36, 2.13)	no				
Vambergue, 2000	21/131	8/108	8.6%	2.16 (1.00, 4.69)	yes	2.5 (1.16, 5.4)	yes	Pre-pregnancy, BMI > 27, maternal age >35, multiparity, educational level.	
Total (95% CI)	**843**	**15,220**	**100.0%**	**1.44 (1.13, 1.82)**					**Adding adjusted estimates would not change significance of overall result**
CC false-positive vs. no GDM									
Hillier, 2007	122/999	905/7,609	43.8%	1.03 (0.86, 1.23)	no	n/a	n/a		
Lapolla, 2007	8/128	8/334	3.8%	2.61 (1.00, 6.81)	no	n/a	n/a		
Naylor, 1996	80/580	395/2,940	35.9%	1.03 (0.82, 1.28)	no				
Ricart, 2005	131/1838	21/263	14.9%	0.89 (0.57, 1.39)	yes	1.33 (1.04, 1.72)	yes	Maternal BMI, fetal sex (male), gestational age, maternal age, macrosomia (yes), PIH (yes)	
Schwartz, 1999	2/49	12/112	1.7%	0.38 (0.09, 1.64)	no	n/a	n/a		
Total (95% CI)	**3,594**	**11,258**	**100.0%**	**1.02 (0.85, 1.24)**					**Adding adjusted estimates would not change significance of overall result**
CC 1 abnormal OGTT vs. false-positive									
Hillier, 2007	40/288	122/999	51.7%	1.14 (0.82, 1.59)	no	n/a	n/a		
Kwik, 2007	42/213	19/197	37.8%	2.04 (1.23, 3.39)	no	n/a	n/a		
Lapolla, 2007	3/48	8/128	10.6%	1.00 (0.28, 3.61)	no	n/a	n/a		

Total (95% CI)	**549**	**1,324**	**100.0%**	**1.40 (0.89, 2.20)**					**No change**
NDDG vs no GDM									
Adams, 1998	7/16	5/64	100.0%	5.60 (2.04, 15.35)	no				**No change**
NDDG false positive vs no GDM									
Chico, 2005	15/187	33/529	21.6%	1.29 (0.71, 2.31)	no	n/a	n/a		
Hillier, 2007	27/326	83/1,432	42.9%	1.43 (0.94, 2.17)	no	n/a	n/a		
Retnakaran, 2008	18/128	6/74	9.7%	1.73 (0.72, 4.18)	no	n/a	n/a		
Stamilio, 2004	14/164	95/1,661	25.8%	1.49 (0.87, 2.56)	yes	1.79 (0.91, 3.51)	no	BMI, parity, gestational age at delivery, chronic hypertension, tobacco use, race, midtrimester serum B-fetoprotein and human chorionic gonadotropin levels, maternal age, history of preeclampsia in previous pregnancy	
Total (95% CI)	**805**	**3,696**	**100.0%**	**1.44 (1.10, 1.89)**					**Adding adjusted estimate would not change significance of overall result**
WHO GDM vs no GDM									
Shirazian, 2008	1/10	16/532	100.0%	3.33 (0.49, 22.70)	no	1.34 (0.15, 12)	no		**No change**
WHO IGT vs no GDM									
Jensen, 2003	98/289	696/2,596	100.0%	1.26 (1.06, 1.50)	yes	1.5 (1.1, 2.2)	no	Pre-pregnancy BMI, maternal age,parity, smoking, weight gain during pregnancy, gestational age, anamnestic risk indicators for GDM, ethnic background and clinical centre.	**No change**
IADPSG GDM vs no GDM									
Lapolla, 2011	12/112	145/1,815	78.8%	1.34 (0.77, 2.34)	no	n/a	n/a		

Morikawa, 2010	1/43	0/160	21.2%	10.98 (0.46, 264.81)	no	n/a	n/a		
Total (95% CI)	**155**	**1,975**	**100.0%**	**2.09 (0.39, 11.33)**					**No change**
Shoulder dystocia									
CC GDM vs. no GDM									
Cheng, 2009	9/273	237/ 13,940	48.40%	1.94 (1.01, 3.73)	yes	2.24 (1.03,4.88)	no	Parity, maternal age, race or ethnicity, gestational weight gain, gestational age at delivery, year of delivery, epidural anesthesia, induction of labor, (with mode of delivery and episiotomy additionally controlled for perineal laceration, postpartum hemorrhage, shoulder dystocia, and birth trauma)	
Chou, 2010	2/489	11/10,116	9.2%	3.76 (0.84, 16.92)	no	n/a	n/a		
Landon, 2011	18/455	3/423	14.1%	5.58 (1.65, 18.80)	yes	5.44 (1.81, 20.1)	no	Maternal age, gestational age at enrollment and at delivery, parity, BMI, and race and ethnicity	
Langer, 2005	14/555	7/1,110	25.6%	4.00 (1.62, 9.85)	no	n/a	n/a		
Pennison, 2001	1/43	1/69	2.8%	1.60 (0.10, 24.99)	no	n/a	n/a		
Total (95% CI)	**1,815**	**25,658**	**100.0%**	**2.86 (1.81, 4.51)**					**Adding adjusted estimate would not change significance of overall result**
CC GDM vs. false-positive									
Berggren, 2011	24/460	109/3,117	100.0%	1.49 (0.97, 2.30)	yes	1.41 (0.91,2.18)	no	Parity, maternal delivery age over 35 years, ethnicity, delivery year; cesarean and operative deliveries were also controlled for prior cesarean	**No change**
CC 1 abnormal OGTT vs. no GDM									

Vambergue, 2000	1/131	4/108	100.0%	0.20 (0.02, 1.82)	no	n/a	n/a		**No change**
CC 1 abnormal OGTT vs. false-positive									
Kwik, 2007	11/213	2/197	100.0%	5.09 (1.14, 22.66)	no	n/a	n/a		**No change**
NDDG GDM (unrecognized) vs. no GDM									
Adams, 1998	3/16	2/64	100.0%	6.00 (1.09, 32.95)	yes	5.2 (1.1, 30.6)	no	Maternal BMI, age, parity, weight gain, gestational age	**No change**
NDDG false-positive vs. no GDM									
Stamilio, 2004	8/164	29/1,661	100.0%	2.79 (1.30, 6.01)	yes	2.85 (1.25, 6.51)	no	BMI, parity, gestational age at delivery, chronic hypertension, tobacco use, race, midtrimester serum B-fetoprotein and human chorionic gonadotropin levels, maternal age, history of preeclampsia in previous pregnancy	**No change**
WHO IGT vs. no GDM									
Jensen, 2003	8/289	33/2,596	100.0%	2.18 (1.02, 4.67)	yes	1.3 (0.4, 3.9)	yes	Pre-pregnancy BMI, maternal age,parity, smoking, weight gain during pregnancy, gestational age, anamnestic risk indicators for GDM, ethnic background and clinical centre.	**Adjusted estimate not statistically significant**
IADPSG IGT vs. no GDM									
Black, 2010	18/391	268/7,020	100.0%	1.21 (0.76, 1.92)	yes	1.31 (0.80, 2.16)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**
IADPSG IFG vs. no GDM									

Black, 2010	50/886	268/7,020	100.0%	1.48 (1.10, 1.98)	yes	1.45 (1.05, 2.00)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**
IADPSG IGT-2 vs. no GDM									
Black, 2010	5/83	268/7,020	100.0%	1.58 (0.67, 3.72)	yes	1.72 (0.68, 4.35)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**
IADPSG IGT IFG vs. no GDM									
Black, 2010	23/331	268/7,020	100.0%	1.82 (1.21, 2.75)	yes	1.87 (1.18, 2.96)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**
IADPSG IGT vs. IFG									
Black, 2010	18/391	50/886	100.0%	0.82 (0.48, 1.38)	no	n/a	n/a		**No change**
IADPSG IGT vs. IGT-2									
Black, 2010	18/391	5/83	100.0%	0.76 (0.29, 2.00)	no	n/a	n/a		**No change**
IADPSG IGT vs. IGT IFG									
Black, 2010	18/391	23/331	100.0%	0.66 (0.36, 1.21)	no	n/a	n/a		**No change**
IADPSG IFG vs. IGT-2									
Black, 2010	50/886	5/83	100.0%	0.94 (0.38, 2.28)	no	n/a	n/a		**No change**
IADPSG IFT vs. IGT IFG									
Black, 2010	50/886	23/331	100.0%	0.81 (0.50, 1.31)	no	n/a	n/a		**No change**
IADPSG IGT-2 vs. IGT IFG									
Black, 2010	5/83	23/331	100.0%	0.87 (0.34, 2.21)	no	n/a	n/a		**No change**

Fetal birth injury									
NDDG GDM (unrecognized) vs. no GDM									
Adams, 1998	4/16	0/64	100.0%	34.41 (1.95, 608.47)	no	n/a	n/a		**No change**
Neonatal hypoglycemia									
CC GDM vs. No GDM									
Chico, 2005	23/422	202/5,767	35.1%	1.56 (1.02, 2.37)	no	n/a	n/a		
Langer, 2005	100/555	21/1,110	34.8%	9.52 (6.02, 15.08)	no	n/a	n/a		
Pennison, 2001	10/43	5/69	30.1%	3.21 (1.18, 8.76)	no	n/a	n/a		
Total (95% CI)	**1,020**	**6,946**	**100.0%**	**3.64 (0.96, 13.76)**					**No change**
CC GDM vs. 1 abnormal OGTT									
Chico, 2005	23/422	1/59	100.0%	3.22 (0.44, 23.37)	no	n/a	n/a		
CC 1 abnormal OGTT vs. no GDM									
Chico, 2005	1/59	202/5,767	4.0%	0.48 (0.07, 3.39)	no	n/a	n/a		
Corrado, 2009	9/152	26/624	27.8%	1.42 (0.68, 2.97)	no	n/a	n/a		
Rust, 1996	9/78	20/205	27.4%	1.18 (0.56, 2.48)	no	n/a	n/a		
Vambergue, 2000	24/131	14/108	40.0%	1.41 (0.77, 2.60)	no	n/a	n/a		
Total (95% CI)	**420**	**6,704**	**100.0%**	**1.29 (0.88, 1.91)**					**No change**
NDDG GDM vs. No GDM									
Adams, 1998	0/16	0/64	Not estimable	n/a	n/a	n/a	n/a		**No change**
NDDG false-positive vs. no GDM									
Ardawi, 2000	3/187	3/529	100.00%	2.83 (0.58, 13.89)	no	n/a	n/a		**No change**
NDDG 1 abnormal vs. no GDM									
Kim, 2002	2/122	1/577	100.00%	9.60 (0.86, 106.73)	no	n/a	n/a		**No change**
WHO IGT vs. WHO no GDM									
Jensen, 2003	6/281	63/2,596	76.60%	0.88 (0.38, 2.01)	yes	0.7 (0.2, 2.2)	no	Pre-pregnancy BMI, maternal age,parity, smoking, weight gain during pregnancy, gestational age, anamnestic risk indicators for GDM, ethnic background and clinical centre.	
Nord, 1995	2/223	3/391	16.50%	1.17 (0.20, 6.94)	no	n/a	n/a		
Yang, 2002	1/102	1/302	6.90%	2.96 (0.19, 46.91)	no	n/a	n/a		

Total (95% CI)	**606**	**3,289**	**100.00%**	**1.00 (0.49, 2.07)**					**Adding adjusted estimate would not change statistical significance of overall result**
Hyperbilirubinemia									
CC GDM vs. No GDM									
Chico, 2005	17/422	144/5,767	49.80%	1.61 (0.99, 2.64)	no	n/a	n/a		
Langer, 2005	78/555	23/1,110	50.20%	6.78 (4.31, 10.68)	no	n/a	n/a		
Total (95% CI)	**977**	**6,877**	**100.00%**	**3.32 (0.80, 13.74)**					**No change**
CC GDM vs. 1 abnormal OGTT									
Chico, 2005	17422/	1/59	100.00%	2.38 (0.32, 17.53)	no	n/a	n/a		**No change**
CC false-positive vs. no GDM									
Bo, 2004	42/315	4/91	100.00%	3.03 (1.12, 8.23)	no	n/a	n/a		**No change**
CC 1 abnormal OGTT vs. no GDM									
Vambergue, 2000	2/131	0/108	100.00%	4.19 (0.20, 88.20)	no	n/a	n/a		**No change**
NDDG false-positive vs. no GDM									
Ardawi, 2000	22/187	58/529	100.00%	1.07 (0.68, 1.70)	no	n/a	n/a		**No change**
WHO IGT vs. WHO no GDM									
Jensen, 2003	6/281	83/2,596	42.40%	0.67 (0.29, 1.52)	no	n/a	n/a		
Nord, 1995	10/223	28/391	57.60%	0.63 (0.31, 1.26)	no	n/a	n/a		
Total (95% CI)	**504**	**2,987**	**100.00%**	**0.64 (0.38, 1.10)**					**No change**
IADPSG IGT vs. no GDM									
Black, 2010	72/391	980/7,020	100.00%	1.32 (1.06, 1.64)	yes	1.33 (1.02, 1.74)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**
IADPSG IFG vs. no GDM									
Black, 2010	128/886	980/7,020	100.00%	1.03 (0.87, 1.23)	yes	1.04 (0.85, 1.27)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**
IADPSG IGT-2 vs. no GDM									

Black, 2010	18/83	980/7,020	100.00%	1.55 (1.03, 2.35)	yes	1.56 (0.92, 2.65)	yes	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**Adjusted result is not statistically significant**
IADPSG IGT OFG vs. no GDM									
Black, 2010	45/331	980/7,020	100.00%	0.97 (0.74, 1.29)	yes	0.96 (0.69, 1.33)	no	Adjusted for maternal age, race/ethnicity, parity, prepregnancy BMI, gestational weight gain, infant sex, and gestational age at OGTT	**No change**
IADPSG IGT vs. IFG									
Black, 2010	72/391	128/886	100.0%	1.27 (0.98, 1.66)	no	n/a	n/a		**No change**
IADPSG IGT vs. IGT-2									
Black, 2010	72/391	18/83	100.0%	0.85 (0.54, 1.34)	no	n/a	n/a		**No change**
IADPSG IGT vs. IGT IFG									
Black, 2010	72/391	45/331	100.0%	1.35 (0.96, 1.91)	no	n/a	n/a		**No change**
IADPSG IFG vs. IGT-2									
Black, 2010	128/886	18/83	100.0%	0.67 (0.43, 1.03)	no	n/a	n/a		**No change**
IADPSG IFG vs. IGT IFG									
Black, 2010	128/886	45/331	100.0%	1.06 (0.78, 1.46)	no	n/a	n/a		**No change**
IADPSG IGT-2 vs. IGT IFG									
Black, 2010	18/83	45/331	100.0%	1.60 (0.98, 2.61)	no	n/a	n/a		**No change**
Fetal Birth Trauma/ Injury									
CC GDM vs. no GDM									
Cheng, 2009	12/273	516/13,940	100.00%	1.19 (0.68, 2.08)	yes	1.26 (0.66, 2.42)	no	Parity, maternal age, race or ethnicity, gestational weight gain, gestational age at delivery, year of delivery, epidural anesthesia, induction	**No change**

								of labor, (with mode of delivery and episiotomy additionally controlled for perineal laceration, postpartum hemorrhage, shoulder dystocia, and birth trauma)	
NDDG GDM vs. No GDM									
Adams, 1998	4/16	0/64	100.0%	34.41 (1.95, 608.47)	no	n/a	no		**No change**
WHO IGT vs. no GDM									
Nord, 1995	1/223	6/391	100.00%	0.29 (0.04, 2.41)	no	n/a	n/a		
Yang, 2002	0/102	0/302	0.00%	Not estimable	no	n/a	n/a		
Total (95% CI)	**325**	**693**	**100.00%**	**0.29 (0.04, 2.41)**					**No change**
Fetal Morbidity/Mortality									
CC GDM vs. no GDM									
Chico, 2005	0/422	29/5,767	10.10%	0.23 (0.01, 3.78)	no	n/a	n/a		
Chou, 2010	1/489	42/10,116	16.80%	0.49 (0.07, 3.57)	no	n/a	n/a		
Langer, 2005	0/555	0/1,110		Not estimable	no	n/a	n/a		
Lapolla, 2011	18/112	132/1,815	46.80%	2.21 (1.40, 3.48)	no	n/a	n/a		
Ricart, 2005	0/263	25/6350	10.10%	0.47 (0.03, 7.73)	no	n/a	n/a		
Schwartz, 1999	1/154	16/7,207	16.40%	2.92 (0.39, 21.92)	no	n/a	n/a		
Total (95% CI)	**1,995**	**32,365**	**100.00%**	**1.23 (0.46, 3.30)**					**No change**
CC GDM vs. false-positive									
Ricart, 2005	0/263	7/1,838	49.10%	0.46 (0.03, 8.11)	no	n/a	n/a		
Schwartz, 1999	1/154	1/1,066	50.90%	6.92 (0.44, 110.10)	no	n/a	n/a		
Total (95% CI)	**417**	**2,904**	**100.00%**	**1.83 (0.11, 29.41)**					**No change**
CC GDM vs. 1 abnormal OGTT									
Chico, 2005	0/422	0/59	n/a	Not estimable	no	n/a	n/a		**No change**
CC 1 abnormal OGTT vs. no GDM									
Chico, 2005	0/59	29/5,767	3.40%	1.63 (0.10, 26.36)	no	n/a	n/a		
Rust, 1996	15/78	40/205	93.90%	0.99 (0.58, 1.68)	no	n/a	n/a		
Vambergue, 2000	1/131	0/108	2.60%	2.48 (0.10, 60.20)	no	n/a	n/a		
Total (95% CI)	**268**	**6,080**	**100.00%**	**1.03 (0.61, 1.72)**					**No change**
CC false-positive vs. no GDM									

Bo, 2004	4/315	2/91	17.40%	0.58 (0.11, 3.10)	no	n/a	n/a		
Ricart, 2005	7/1,838	25/6,350	70.50%	0.97 (0.42, 2.23)	no	n/a	n/a		
Schwartz, 1999	1/1,066	16/7,207	12.10%	0.42 (0.06, 3.18)	no	n/a	n/a		
Total (95% CI)	**3,219**	**13,648**	**100.00%**	**0.80 (0.40, 1.61)**					**No change**
CC false-positive vs. 1 abnormal OGTT									
Kwik, 2007	0/197	0/213	n/a	Not estimable	no	n/a	n/a		**No change**
NDDG false-positive vs. no GDM									
Ardawi, 2000	2/187	4/529	47.00%	1.41 (0.26, 7.66)	no	n/a	n/a		
Stamilio, 2004	2/164	6/1,661	53.00%	3.38 (0.69, 16.59)	yes	4.61 (0.77, 27.48)	no	BMI, parity, gestational age at delivery, chronic hypertension, tobacco use, race, midtrimester serum B-fetoprotein and human chorionic gonadotropin levels, maternal age, history of preeclampsia in previous pregnancy	
Total (95% CI)	**351**	**2,190**	**100.00%**	**2.24 (0.70, 7.14)**					**Adding adjusted estimate would not change significance of overall result**
NDDG 1 abnormal OGTT vs. no GDM									
Kim, 2002	0/122	2/577	100.00%	0.94 (0.04, 19.69)	no	n/a	n/a		**No change**
WHO IGT vs. no GDM									
Aberg, 2001	1/126	13/4,515	22.90%	2.76 (0.36, 20.91)	no	n/a	n/a		
Nord, 1995	3/223	7/391	52.20%	0.75 (0.20, 2.88)	no	n/a	n/a		
Yang, 2002	2/102	2/302	24.80%	2.96 (0.42, 20.75)	no	n/a	n/a		
Total	**451**	**5,208**	**100.0%**	**1.42 (0.54, 3.75)**					**No change**
IADPSG GDM vs. no GDM									
Lapolla, 2011	18/112	132/1,815	100.00%	2.21 (1.40, 3.48)	no	n/a	n/a		**No change**
Prevalence of Childhood Obesity (>85th percentile)									
CC GDM vs. no GDM									
Hillier, 2007	60/173	1788/7,609	100.00%	1.48 (1.20, 1.82)	yes	1.89 (1.30, 2.76)	no	Maternal age, parity, weight gain during pregnancy, ethnicity, macrosomia at birth	**No change**

								(4,000 g), and sex of child	
CC GDM vs. false-positive									
Hillier, 2007	60/173	233/999	100.00%	1.49 (1.18, 1.88)	no	n/a	n/a		**No change**
CC GDM vs. 1 abnormal OGTT									
Hillier, 2007	60/173	77/288	100.00%	1.30 (0.98, 1.72)	no	n/a	n/a		**No change**
CC false-positive vs. no GDM									
Hillier, 2007	233/999	1788/7,609	100.00%	0.99 (0.88, 1.12)	yes	0.98 (0.81, 1.17)	no	Maternal age, parity, weight gain during pregnancy, ethnicity, macrosomia at birth (4,000 g), and sex of child	**No change**
CC false-positive vs. 1 abnormal OGTT									
Hillier, 2007	233/999	77/288	100.00%	0.87 (0.70, 1.09)	no	n/a	n/a		**No change**
CC 1 abnormal OGTT vs. no GDM									
Hillier, 2007	77/288	1788/7,609	100.00%	1.14 (0.94, 1.38)	yes	1.37 (1.01, 1.84)	yes (result becomes significant)	Maternal age, parity, weight gain during pregnancy, ethnicity, macrosomia at birth (4,000 g), and sex of child	

* The information presented in these columns is number of patients with the outcome / numbers of patients per group.
† The effect estimates are risk ratios with 95% confidence intervals.
BMI = body mass index; CC = Carpenter-Coustan; CI = confidence interval; GDM = gestational diabetes mellitus; IADPSG = International Association of Diabetes and Pregnancy Study Groups; IFG = impaired fasting glucose; IFT = impaired fasting tolerance; IGT = impaired glucose tolerance; IGT-2 = double impaired glucose tolerance; NDDG = National Diabetes Data Group; n = number of patients with the outcome; N = numbers of patients per group; n/a = not applicable; OGTT = oral glucose tolerance test; PIH = Pregnancy induced hypertension;SD = standard deviation; WHO = World Health Organization