Screening for Gestational Diabetes
Updated Evidence Report and Systematic Review for the US Preventive Services Task Force

Jennifer Pillay, MSc; Lois Donovan, MD; Samantha Guitard, MSc; Bernadette Zakher, MD; Michelle Gates, PhD; Allison Gates, PhD; Ben Vandermeer, MSc; Christina Bougatsos, MPH; Roger Chou, MD; Lisa Hartling, PhD

IMPORTANCE Gestational diabetes is associated with several poor health outcomes.

OBJECTIVE To update the 2012 review on screening for gestational diabetes to inform the US Preventive Services Task Force.

DATA SOURCES MEDLINE, EMBASE, and CINAHL (2010 to May 2020), ClinicalTrials.gov, reference lists; surveillance through June 2021.

STUDY SELECTION English-language intervention studies for screening and treatment; observational studies on screening; prospective studies on screening test accuracy.

DATA EXTRACTION AND SYNTHESIS Dual review of titles/abstracts, full-text articles, and study quality. Single-reviewer data abstraction with verification. Random-effects meta-analysis or bivariate analysis (accuracy).

MAIN OUTCOMES AND MEASURES Pregnancy, fetal/neonatal, and long-term health outcomes; harms of screening; accuracy.

RESULTS A total of 76 studies were included (18 randomized clinical trials [RCTs] [n = 31 241], 2 nonrandomized intervention studies [n = 190], 56 observational studies [n = 261 678]). Direct evidence on benefits of screening vs no screening was limited to 4 observational studies with inconsistent findings and methodological limitations. Screening was not significantly associated with serious or long-term harm. In 5 RCTs (n = 25 772), 1-step (International Association of Diabetes and Pregnancy Study Group) vs 2-step (Carpenter and Coustan) screening was significantly associated with increased likelihood of gestational diabetes (11.5% vs 4.9%) but no improved health outcomes. At or after 24 weeks of gestation, oral glucose challenge tests with 140- and 135-mg/dL cutoffs had sensitivities of 82% and 93%, respectively, and specificities of 82% and 79%, respectively, against Carpenter and Coustan criteria, and a test with a 140-mg/dL cutoff had sensitivity of 85% and specificity of 81% against the National Diabetes Group Data criteria. Fasting plasma glucose tests with cutoffs of 85 and 90 mg/dL had sensitivities of 88% and 81% and specificities of 73% and 82%, respectively, against Carpenter and Coustan criteria. Based on 8 RCTs and 1 nonrandomized study (n = 3982), treatment was significantly associated with decreased risk of primary cesarean deliveries (relative risk [RR], 0.70 [95% CI, 0.54-0.91]; absolute risk difference [ARD], 5.3%), shoulder dystocia (RR, 0.42 [95% CI, 0.23-0.77]; ARD, 1.3%), macrosomia (RR, 0.53 [95% CI, 0.41-0.68]; ARD, 8.9%), large for gestational age (RR, 0.56 [95% CI, 0.47-0.66]; ARD, 8.4%), birth injuries (odds ratio, 0.33 [95% CI, 0.11-0.99]; ARD, 0.2%), and neonatal intensive care unit admissions (RR, 0.73 [95% CI, 0.53-0.99]; ARD, 2.0%). The association with reduction in preterm deliveries was not significant (RR, 0.75 [95% CI, 0.56-1.01]).

CONCLUSIONS AND RELEVANCE Direct evidence on screening vs no screening remains limited. One- vs 2-step screening was not significantly associated with improved health outcomes. At or after 24 weeks of gestation, treatment of gestational diabetes was significantly associated with improved health outcomes.
Gestational diabetes is diabetes that develops during pregnancy. The prevalence of gestational diabetes in the US has typically been estimated at 5.6% to 9.2% when measured from 2007 to 2016 but may be up to 3-fold higher depending on the diagnostic criteria used. Gestational diabetes is usually asymptomatic but is associated with increased risk for several pregnancy and neonatal complications.

In 2014, the US Preventive Services Task Force (USPSTF) recommended screening for gestational diabetes in asymptomatic pregnant women after 24 weeks of gestation (B recommendation). The USPSTF found that evidence was insufficient to screen before 24 weeks of gestation (I statement). This evidence report was conducted to update the 2012 review to inform updated USPSTF recommendations.

Methods

Scope of the Review
Detailed methods and additional study details are available in the full evidence report. Figure 1 shows the analytic framework and key questions (KQs) that guided the review. KQ5 is addressed only in the full report. KQ3, comparing different screening strategies, was added for this update. This review did not address screening for preexisting or overt diabetes in early pregnancy.

Data Sources and Searches
Ovid MEDLINE and EMBASE, and CINAHL via EBSCOhost, were searched from 2010 to May 22, 2020 (eMethods 1 in the Supplement). Clinical trial registries and reference lists (including the 2012 review) were reviewed. Ongoing surveillance was conducted to identify major studies published since May 2020 that may affect the conclusions or understanding of the evidence and the related USPSTF recommendation.

Study Selection
Two investigators independently reviewed titles and abstracts, then full-text articles using predefined eligibility criteria (eMethods 2 in the Supplement). The population for screening and test accuracy was pregnant women without known preexisting diabetes mellitus. For treatment, the population was women with gestational diabetes or hyperglycemia. For harms and harms of screening, comparative effectiveness of screening approaches, and screening test accuracy, studies using 1-step (diagnostic test only) or 2-step (diagnostic test in women with a positive screening test result) screening strategies at any time during pregnancy were included (eMethods 3 in the Supplement). In 2-step strategies, the screening test was measurement of fasting plasma glucose level, a 50-g oral glucose challenge test (OGCT), a risk factor-based tool, or glycated hemoglobin (HbA1c) concentration. For benefits of screening and treatment, comparisons were against no screening or treatment, respectively. For harms of screening, studies comparing outcomes before and after a gestational diabetes diagnosis or comparing women with gestational diabetes aware of their diagnosis vs those unaware were included. To evaluate potential labeling harms, studies on receipt of delivery and perinatal interventions among women diagnosed with gestational diabetes vs those without a diagnosis were included. For accuracy, the reference standard was a currently recommended oral glucose tolerance test (OGTT), mainly using Carpenter and Coustan, the National Diabetes Data Group, or the International Association of Diabetes and Pregnancy Study Group (IADPSG) Consensus Panel diagnostic criteria. Intermediate and health outcomes are listed in Figure 1. Studies had to be published on or after 1995 and conducted in settings applicable to primary care.

Randomized clinical trials (RCTs) and nonrandomized controlled intervention studies were included for screening and treatment; for screening vs no screening, controlled observational studies were also included because of anticipated lack of intervention studies and to assess potential harms. For screening test accuracy, prospective cohort studies in which at least a sample of screen-negative women underwent the reference standard were included. Studies on risk factor strategies or models had to examine a validation cohort.

Data Extraction and Quality Assessment
One reviewer abstracted data from the studies; a second reviewer verified accuracy and completeness. Outcomes related to hypertension in pregnancy were classified as preeclampsia, gestational hypertensive disorders in pregnancy (mixed). For cesarean delivery, primary (first) cesarean deliveries were prioritized, but total and emergency cesarean rates were also evaluated. Two reviewers independently assessed the methodological quality of eligible studies using design-specific tools (eMethods 4 in the Supplement). Disagreements were resolved by consensus and, if necessary, consultation with a third reviewer. Studies were rated as “good,” “fair,” or “poor,” based on the seriousness of methodological shortcomings.

Data Synthesis and Analysis
For intervention effects using relative risks (RRs), meta-analyses used random-effects models in Review Manager version 5.1 (The Cochrane Collaboration). When moderate or greater statistical heterogeneity ($I^2 \geq 40\%$) was observed, sensitivity analysis was performed using the profile likelihood method in Stata version 14.2 (StataCorp); these analyses did not change any of the conclusions, but results are available in the full report (and for KQ3 are reported in eTables 3 and 4 in the Supplement). Pooled absolute risk differences (ARDs) were calculated when RRs were statistically significant and for all analyses with at least 1 zero-event study. Heterogeneity was explored with sensitivity analyses using predefined variables (eg, study quality, setting, differing outcome definitions); findings of within-study subgroup analyses were extracted.

For diagnostic accuracy, analyses were stratified by timing of the index test in pregnancy and comparison, including different test thresholds. If more than 3 studies were included for a particular comparison, sensitivities and specificities were pooled using bivariate analysis (metandi program in Stata version 14.2) with construction of hierarchical summary receiver operator characteristic curves. The aggregate strength of evidence was assessed for each outcome, using the Agency for Healthcare Research and Quality methods guidance, based on the number, quality, and size of studies and the consistency and precision of results between studies. Significance testing was 2-tailed; $P < .05$ was considered statistically significant.
Figure 1. Analytic Framework: Screening for Gestational Diabetes

Key questions

1. Does screening for gestational diabetes reduce poor health outcomes?
   a. Does screening for gestational diabetes reduce poor intermediate outcomes?
   b. Does the effectiveness of screening for gestational diabetes vary according to maternal subgroup characteristics, including timing during pregnancy, previous gestational diabetes diagnosis, family history of type 2 diabetes, body mass index, age, or race/ethnicity?

2. What are the harms of screening for and diagnosis of gestational diabetes to the mother, fetus, or neonate?

3. What is the comparative effectiveness of different screening strategies for gestational diabetes on health outcomes?
   a. What is the comparative effectiveness of different screening strategies for gestational diabetes on intermediate outcomes?
   b. Does the comparative effectiveness of different screening strategies for gestational diabetes vary according to maternal subgroup characteristics, including timing during pregnancy, previous gestational diabetes diagnosis, family history of type 2 diabetes, body mass index, age, or race/ethnicity?

4. What is the diagnostic accuracy of commonly used screening tests for gestational diabetes?
   a. What is the diagnostic accuracy of commonly used screening tests for gestational diabetes on health outcomes?
   b. Does the accuracy of commonly used screening tests for gestational diabetes vary according to maternal subgroup characteristics, including timing during pregnancy, body mass index, age, race/ethnicity, or prevalence of gestational diabetes?

5. What is the association between diagnosis of gestational diabetes and outcomes in women meeting more inclusive but not less inclusive diagnostic criteria for gestational diabetes? (see full report for details)

6. Does treatment of gestational diabetes during pregnancy reduce poor health outcomes?
   a. Does treatment of gestational diabetes during pregnancy reduce poor intermediate outcomes?
   b. Does the effectiveness of treatment of gestational diabetes vary according to maternal subgroup characteristics, including timing and criteria used for diagnosis during pregnancy, severity of hyperglycemia, body mass index, age, or race/ethnicity?

7. What are the harms of treatment of gestational diabetes, including severe maternal and fetal/neonatal hypoglycemia, delivery of neonates who are small for gestational age, and poor long-term growth and development outcomes in the child?

See the USPSTF Procedure Manual for information about analytic frameworks. KQ indicates key question. *No assumptions will be made about whether hyperglycemia discovered early in pregnancy (e.g., in first trimester) is gestational diabetes or some other form; gestational diabetes will include all women with hyperglycemia but not meeting criteria for overt diabetes at any time during pregnancy. Using 2-step (screening first and, when indicated, diagnostic tests second) or 1-step (diagnostic tests only) strategies, each based on various criteria and thresholds, and offering treatment to patients diagnosed with gestational diabetes.
Results

Across the KQs reported in this manuscript, a total of 76 studies were included (18 RCTs [different screening strategies, n = 27 19624-24; treatment benefits and harms, n = 40452-25], 2 nonrandomized controlled intervention studies of treatment [n = 190]36-37 and 56 observational studies [screening benefits, n = 433628-41, screening harms, n = 166 08222-28, diagnostic accuracy, n = 91 26049-93]) (Figure 2). Twenty-three studies were carried over from the prior USPSTF review and 52 new studies were added.

Benefits and Harms of Screening and Different Screening Strategies

Key Question 1. Does screening for gestational diabetes reduce poor health or intermediate outcomes? Does effectiveness vary according to prespecified subgroups?

No RCTs or nonrandomized intervention studies were identified. Four retrospective observational studies (n = 433628-41 compared women who underwent screening for gestational diabetes with women who were not screened (eTable 1 in the Supplement). Two studies28,40 from the previous USPSTF review evaluating selected women showed no significant effect of screening; however, sample sizes were small and estimates imprecise. One new study (n = 1012)21 found 1-step screening of at-risk women was significantly associated with a reduction in late (≥28 weeks of gestation) stillbirth (adjusted odds ratio [OR], 0.68 [95% CI, 0.47 to 0.97]). The other new study (n = 2780)29 found universal 2-step screening, with early screening offered to women with risk factors, significantly associated with fewer cesarean deliveries (RR, 0.78 [95% CI, 0.66 to 0.92]; ARD, -4.8% [95% CI, -8.2% to -1.5%], birth injuries (fracture or dislocation; RR, 0.47 [95% CI, 0.23 to 0.97]; ARD, -0.9% [95% CI, -1.9% to 0.1%]), and admissions to the neonatal intensive care unit (NICU) (RR, 0.67 [95% CI, 0.58 to 0.78]; ARD, -8.7% [95% CI, -12.3% to -5.2%]) compared with historical controls. Prespecified analyses found screening in first trimester was significantly associated with decreased likelihood of NICU admissions vs second-trimester screening but with no significant difference for other outcomes. Both new studies were susceptible to confounding and selection bias.

Key Question 2. What are the harms of screening for and diagnosis of gestational diabetes to the mother, fetus, or neonate?

All 7 studies32-48 identified for KQ2 were new to this update (eTable 2 in the Supplement). No significant differences were found in anxiety and depressive symptoms before and after screening for those with negative or false-positive results in 2 cohort studies (n = 1015).44-48 One study (n = 100)42 found that anxiety symptoms scores were slightly higher (6 points on 60-point scale; P = .007) for women with vs without gestational diabetes immediately after receiving results but not significantly higher at gestational week 36 or 6 weeks postpartum.

One good-quality cohort study (n = 3778)46 found that the association between macrosomia and cesarean delivery in women with normoglycemia or untreated borderline gestational diabetes was not observed in those with treated gestational diabetes, suggesting that a gestational diabetes diagnosis may have increased the propensity to perform cesarean deliveries. Three large US studies (n = 161 182)43,45,47 found some differences in hospital experiences (eg, adjusted OR, 0.55 [95% CI, 0.36 to 0.85] for fewer newborns staying in mother’s room) potentially related to labeling because of a gestational diabetes diagnosis. However, there were unmeasured potentially confounding factors such as rates of neonatal hypoglycemia, breastfeeding intentions, and varying hospital policies.

Key Question 3. What is the comparative effectiveness of different screening strategies for gestational diabetes on health and intermediate outcomes? Does comparative effectiveness vary according to prespecified subgroups?

IADPSG vs Carpenter and Coustan Screening

Five RCTs (n = 25 77223-24) examined universal screening at 24 to 28 weeks of gestation with the 1-step IADPSG vs 2-step Carpenter and Coustan criteria (Table 1). Three trials were rated fair quality and 222,24 good quality. In the largest trial (n = 23 792),20 25% of women allocated to 1-step screening crossed over to 2-step screening, although results remained similar in an intention-to-treat analysis adjusted for gestational diabetes and adherence. Of the women in this trial’s 2-step group, 1.4% received treatment despite having no diagnosis (only an isolated fasting glucose level ≥95 mg/dL), but the authors’ sensitivity analysis for the outcome of large for gestational age showed no evidence that this reclassification affected results. Data from another trial (n = 786)23 were obtained from a systematic review24 and could not be verified.

One-step vs 2-step screening was significantly associated with identification of gestational diabetes in 11.5% vs 4.9% of participants but was not significantly associated with differences in any pregnancy or fetal/neonatal outcome (eTables 3 and 4, eFigures 1-3 in the Supplement). There was statistical heterogeneity in some analyses in which a fair-quality trial23 found significant associations favoring 1-step screening, whereas findings between 1 good-quality trial24 and the largest trial (fair quality)20 were similar. In the largest trial, 1-step screening significantly increased risk for neonatal hypoglycemia vs 2-step screening, although this may have been in part due to the routine surveillance of neonates with risk factors including diagnosis of maternal gestational diabetes (eFigure 2 in the Supplement). In 1 trial (n = 921)24 in which all women randomized to 2-step screening underwent the 100-g OGTT (to assist with blind), 2-step screening was associated with significantly more testing-related adverse events than 1-step screening (eg, reactive hypoglycemia, vomiting, nausea). However, these findings overestimated harms of 2-step screening in clinical practice, in which only women with an abnormal 50-g OGCT result would undergo the 100-g OGTT.

Early vs Usual Timing for Carpenter and Coustan Screening

One good-quality RCT (n = 922)29 enrolling obese women found early (14 to 20 weeks) vs usual timing of screening with Carpenter and Coustan criteria potentially associated with increased risk of preclampsia, but the difference was not statistically significant (RR, 1.42 [95% CI, 0.99 to 2.05]; ARD, 4.0% [95% CI, 0.0% to 8.0%]). There were no significant differences for other outcomes, although some estimates were imprecise (eTables 3 and 4 in the Supplement).

Diagnostic Test Accuracy

Key Question 4. What is the diagnostic accuracy of commonly used screening tests for gestational diabetes? Does accuracy vary according to maternal subgroup characteristics?
Figure 2. Literature Search Flow Diagram: Screening for Gestational Diabetes

- 98 Potentially relevant unique citations from prior USPSTF review
- 12,185 Potentially relevant unique citations identified through database searches
- 21 Potentially relevant unique citations identified from reference lists of included studies and reviews or surveillance

- 11,244 Citations excluded at title and abstract stage
- 1060 Full-text articles reviewed for eligibility or KQs and contextual questions

- 942 Articles pulled or excluded
  - 46 Pulled for contextual questions only
  - 896 Excluded
    - 211 Wrong study design
    - 121 Wrong comparison (other)
    - 121 Wrong comparison (other)
    - 100 Wrong population
    - 90 Wrong index test (KQ4)
    - 61 Not primary research
    - 48 Wrong criteria (KQ5)
    - 40 Wrong intervention
    - 40 Wrong comparison (KQ2)
    - 30 Wrong population (development cohort for KQ4)
    - 14 Duplicate
    - 8 Protocol
    - 4 Non-English language
    - 4 Other

- 118 Articles (107 studies) included

- 4 Studies included for KQ1
  - 2 Prior review
  - 2 Update
- 7 Studies included for KQ2
  - 0 Prior review
  - 7 Update
- 7 Studies included for KQ3
  - 0 Prior review
  - 7 Update
- 45 Studies included for KQ4
  - 16 Prior review
  - 29 Update
- 31 Studies included for KQ5
  - 13 Prior review
  - 18 Update
- 14 Studies included for KQ6
  - 5 Prior review
  - 9 Update
- 14 Studies included for KQ7
  - 5 Prior review
  - 9 Update

* Forty-one studies included in the prior US Preventive Services Task Force (USPSTF) review were excluded for this update, largely because they were observational studies on treatment (n = 6) and studies on accuracy (n = 35) that did not use a predefined reference standard or did not test at least a sample of women with the reference standard. All details for key question (KQ) 5 are in the full report.9
| Source, country | Study design (No. enrolled; No. analyzed) | Quality | Inclusion criteria | Exclusion criteria | Screening strategy | Treatment differences | Gestational wk at delivery, mean (SD), wGA |
|----------------|-----------------------------------------|--------|------------------|-------------------|-------------------|----------------------|----------------------------------------|
| Davis et al,24 2021 US | RCT (n = 921; 855 analyzed) | Good | Women aged 18-45 y and at 18-28 wk, 6 d gestational age receiving care at 1 of 10 obstetric clinics | Preexisting diabetes (glucose ≥200 mg/dL [≥11.1 mmol/L] on OGCT during baseline visit), diabetes diagnosed before 24 wGA, multifetal gestation, hypertension requiring medications, any corticosteroid use 30 d before enrollment, major congenital anomaly, anticipated preterm delivery before 28 wGA, inability to complete glucose testing before 30 wGA, HIV infection, liver disease, and history of gastric bypass surgery or other conditions that precluded glucose consumption for OGTT | IADPSG (universal, 75 g 1-step) at 25-32 wGA (n = 461; 14.4% with gestational diabetes) | Carpenter and Coustan (universal, 100-g 2-step; OGCT ≥130 mg/dL) with OGTT at 25-32 wGA (n = 460; 4.5% with gestational diabetes) | Gestational diabetes treatment occurred per routine clinical care; individualized nutritional counseling by CDE in group or individual setting, SMBG, medical management as per treating physician; Medication use among participants, group 1 vs group 2: 9.3% vs 2.4% |
| Hillier et al,20 2021 US | RCT (n = 35 579 randomized at first prenatal visit; 23 792 analyzed; see exclusion criteria) | Fair (open-label and high crossover, but adjusted results very similar) | All pregnant women aged ≥18 y receiving care at ≥2 large health maintenance organizations | Preexisting diabetes (before randomization); postrandomization exclusions of 33.1% (of 35 579) mainly due to miscarriage (31.8%) but also multiple gestation, aged <18 y, previous bariatric surgery, and change in insurance Baseline characteristics very similar between groups | IADPSG (universal, 75 g 1-step) at 24-28 wGA, or in first trimester if obese or high-risk (criteria NR; 10% used HbA1c or FPG) (n = 11 922; 1967 [16.5%] with gestational diabetes) 25% received Carpenter and Coustan as diagnostic test | Carpenter and Coustan (universal, 100-g 2-step; OGCT ≥130 or 140 mg/dL at 24-28 wGA, or in first trimester if obese or high-risk (criteria NR; 9% used HbA1c or FPG) (n = 11 870; 1009 [8.5%] with gestational diabetes) | Same treatment protocol between groups; referred to a dietitian for individually tailored diet and lifestyle recommendations, and SMBG, with medication (90% insulin) added when targets not met Insulin/medication among those with gestational diabetes, group 1 vs group 2: 42.6% vs 45.6%; Carpenter and Coustan women (n = 162) with isolated FPG ≥95 mg/dL on OGTT received treatment but were not diagnosed with gestational diabetes; sensitivity analysis for large for gestational age showed no evidence of effect from reclassifying these women as having gestational diabetes |
| Khalifeh et al,21 2020 US | RCT (n = 284; 226 analyzed) | Fair (open-label; 79% women analyzed) | Women without preexisting diabetes | Women with history of preexisting diabetes or a history of bariatric surgery; failure to attend screening (after randomization; n = 35) | IADPSG (universal, 75 g 1-step) at 24-28 wGA or at initial prenatal visit if ≥1 risk factor (and repeated at 24-28 wGA if negative) (n = 123; 10 [8.1%] with gestational diabetes) | Carpenter and Coustan (universal, 100-g 2-step; OGCT ≥135 mg/dL) at 24-28 wGA, or at initial prenatal visit if ≥1 risk factor (and repeated at 24-28 wGA if negative) (n = 126; 7 [5.6%] with gestational diabetes) | Treatment for gestational diabetes was the same regardless of group allocation; delivery at 39 wk to 39 wk, 6 d gestation was recommended to all women with gestational diabetes; use of medication or insulin, group 1 vs group 2: 4.1% vs 3.2% |

(continued)
| Source, country | Study design | Treatment differences | Gestational wk at delivery, mean (SD) | Treatment for gestational diabetes | Inclusion criteria | Exclusion criteria | Screening strategy | Treatment differences |
|----------------|-------------|-----------------------|--------------------------------------|----------------------------------|-------------------|------------------|-------------------|----------------------|
| **Scifres et al, 2015** | US | RCT (n = 47; 47 analyzed) | Group 1: 39.3 (1.1) Group 2: 39.6 (1.3) | Treatment for gestational diabetes performed according to clinical care standards of each participant's primary care provider | Aged 18-45 y, singleton pregnancy between 18-24 wGA receiving prenatal care at an outpatient obstetric clinic at a large academic teaching hospital | Preexisting diabetes or positive screen for diabetes within first trimester (<24 wGA), multiple gestations, corticosteroid use 30 d prior to delivery, inability to complete testing before 30 completed wGA, or anticipated preterm delivery for maternal or fetal indications | IADPSG (universal, 75-g 1-step) at 24-28 wGA (n = 24; 1 [4%] with gestational diabetes) Carpenter and Coustan (universal, 100-g 2-step; OGCT ≥130 mg/dL) at 24-28 wGA (n = 23; 0 with gestational diabetes) | Initial OGCT; if >200 mg/dL, excluded and not randomized |
| **Sevket et al, 2014** | Turkey | RCT (n = 856; 786 analyzed) | NR | Treatment for gestational diabetes was the same regardless of group allocation; endocrinologist care with SMBG, diet, and, if needed, medication | Women 24-28 wGA, referred for gestational diabetes screening and coming for screening visit | Preexisting diabetes, major medical illness (cardiac disease, HIV, hemoglobinopathy, oxygen requirement), bariatric surgery, prior cesarean delivery, known fetal anomalies, chronic prednisone use | IADPSG (universal, 75-g 1-step) at 24-28 wGA (n = 386; 56 [14.5%] with gestational diabetes) Carpenter and Coustan (universal, 100-g 2-step; OGCT ≥140 mg/dL) at 24-28 wGA (n = 400; 24 [6%] with gestational diabetes) | All patients first given OGCT and, if >200 mg/dL, excluded and not randomized |
| **Harper et al, 2020** | US | RCT (n = 962; 922 analyzed) | NR | Treatment for gestational diabetes was the same regardless of group allocation (diabetes educator and SMBG; insulin, glyburide, or metformin chosen at discretion of primary care provider if glucose targets not met); use of insulin, group 1 vs group 2: 2.4% vs 0.7%, *P* = .03; use of any diabetes medication, group 1 vs group 2: 6.8% vs 4.3%, *P* = .34 | Obese (BMI ≥30), nonanomalous, singleton gestations, receiving prenatal care at <20 wGA at the university hospital | Preexisting diabetes, history of macrosomic infant (>4 kg), or polycystic ovarian syndrome | Early screening by Carpenter and Coustan (universal, 100-g 2-step; OGCT ≥135 mg/dL) at 14-20 wGA If negative, underwent repeat screening at 24-28 wGA (n = 454; 69 [17.8%] with gestational diabetes) All had HbA1c at 14-20 wGA and 24-28 wGA; >6.5% diagnostic for gestational diabetes; if HbA1c 6.2%-6.5%, underwent 2-step screening for gestational diabetes 84.3% received early screening | Routine screening by Carpenter and Coustan (universal, 100-g 2-step; OGCT ≥135 mg/dL) at 24-28 wGA (n = 458; 56 [12.6%] with gestational diabetes; 1 with gestational diabetes before 24 wk) All had HbA1c at 14-20 wGA and 24-28 wGA; >6.5% diagnostic for gestational diabetes; if HbA1c 6.2%-6.5%, underwent 2-step screening for gestational diabetes 95.9% received screening |
Across 45 prospective cohort studies on diagnostic accuracy, mean sample size was 500 (range, 42-24 854), mean age was 28.8 years (range, 25-32.7), and mean body mass index (BMI, calculated as weight in kilograms divided by height in meters squared) from 22 studies was 24.6 (range, 21.1-28.1). Studies were conducted in 25 countries. Seventeen studies (38%) were rated good quality and 28 (62%) fair quality. No study reported how accuracy varied according to patient characteristics.

50-g Oral Glucose Challenge Test
eFigure 4 in the Supplement shows findings for the OGCT; results from pooled analyses are summarized in eTable 5 in the Supplement. Against Carpenter and Coustan criteria, at a 140-mg/dL cutoff, the pooled sensitivity and specificity (8 studies, n = 6190)53,60,69,73,76,81,83,90 were 81.9% (95% CI, 68.3% to 90.4%) and 81.8% (95% CI, 71.2% to 89.1%), respectively. Against the National Diabetes Data Group criteria (6 studies, n = 5375),58,66,67,72,81,88 the 140-mg/dL cutoff had a sensitivity of 85% (95% CI, 72.0% to 92.6%) and specificity of 81.2% (95% CI, 75.9% to 85.6%). Using a 135-mg/dL cutoff (4 studies, n = 1554),50,73,76,83 resulted in higher sensitivity (93.3% [95% CI, 23.7% to 99.8%]), reference, Carpenter and Coustan criteria but lower specificity (78.9% [95% CI, 53.3% to 92.5%]) than the 140-mg/dL cutoff. At a cutoff of 130 mg/dL, findings against Carpenter and Coustan criteria were inconsistent from 3 studies (n = 1034) (eFigure 4 in the Supplement).50,76,83 Sensitivity of the OGCT against IADPSG criteria was low across all cutoffs in 2 good-quality studies (n = 2091; eFigure 4 in the Supplement).50,76,83 Sensitivity of the OGCT against IADPSG criteria were inconsistent from 3 studies (n = 430)51 and National Diabetes Data Group (1 study; n = 114)85 or IADPSG (n = 258)93 criteria; however, specificity was highly variable (17% to 80%).

Benefits and Harms of Treatment

Key Question 6. Does treatment of gestational diabetes during pregnancy reduce poor health and intermediate outcomes? Does effectiveness vary according to maternal subgroup characteristics?

Eleven RCTs (n = 4045)25-35 and 2 nonrandomized controlled intervention studies (n = 190)36,37 addressed treatment of gestational diabetes. Mean sample size was 326 (range, 21-1000), mean age was 29.2 years (range, 26.3-32.6), and mean BMI was 28.4 (range, 22.9-34.5) (Table 2). Four studies27,38,32,34 were rated good quality and the others fair quality.

Treatment at 24 to 28 Weeks of Gestation

Like the prior USPSTF review, 2 large good-quality RCTs (n = 1958)27,32 contributed a substantial proportion (40%-90%) of the events for many analyses. Four new studies were added28,31,35,36 and 6 new publications95-100 for 1 large previously included trial32 provided data for long-term outcomes or subgroup analyses. Based on trial inclusion criteria, findings are most applicable to adult women identified using 2-step screening, though there were some differences across trials in eligibility criteria, baseline glycermia, and treatment protocols (Table 2). Apart from 1 trial63 that did not report data, weeks of gestation at delivery was similar between groups in all trials.

Treatment of gestational diabetes was significantly associated with lower risk of common cesarean deliveries vs no treatment (3 studies; RR, 0.70 [95% CI, 0.54 to 0.91]; I² = 0%; ARD, −5.3% [95% CI, −10.3% to −0.24%]),25,32,36 Treatment was also associated, though not significantly, with fewer preterm deliveries (4 studies; RR, 0.75 [95% CI, 0.56 to 1.01]; I² = 0%; ARD, −2.6% [95% CI, −4.9% to 0.02%]).31,32,35,36 (Figure 3). There was no significant association but marked inconsistency for pre eclampsia (Figure 4; 6 studies25,28,31,32,35,36) and hypertensive disorders in pregnancy (3 trials27,32,35); findings appeared sensitive to inclusion of a trial35 from a country not rated as “very high” on the Human Development Index (Figure 4). Treatment was not significantly associated with reduced risk of gestational hypertension (2 trials32,35; some impression), total cesarean deliveries (8 trials25-29,31,32,35), emergency cesarean deliveries (1 trial27), induction of labor (5 trials25,28,32,35,36), or maternal birth trauma (2 studies27,36).

For fetal/neonatal outcomes, treatment was significantly associated with lower risk of shoulder dystocia (4 trials; RR, 0.42 [95% CI, 0.23 to 0.77]; I² = 0%; ARD, −1.3% [95% CI, −4.3% to −1.6%]),25,27,32,35 macrosomia (8 studies; RR, 0.53 [95% CI, 0.41 to 0.68]; I² = 42%; ARD, −8.9% [-12.0% to −5.9%]),25,27,29,31,32,35,36 LGA (7 trials; RR, 0.56 [95% CI, 0.47 to 0.66]; I² = 0%; ARD, −8.4% [95% CI, −10.8% to −6.1%]),26,28,31,32,35,36 and NICU admissions (5 trials; RR, 0.73 [95% CI, 0.53 to 0.99]; I² = 0%; ARD, −2.0% [95% CI, −4.5% to 0.5%]).26,28,31,32,35 Treatment was significantly associated with reduced risk of birth injury (eg, fracture or nerve palsies) in 3 trials that reported at least 1 event (OR, 0.33 [95% CI, 0.11 to 0.99]; I² = 0%)27,32,33 but not when including 4 zero-event trials (ARD, −0.2% [95% CI, −0.6% to 0.2%]).27,29,31,32,35,36 There was no significant association between treatment and risk of mortality (6 trials),27-29,32,35,36 respiratory distress syndrome (2 trials),27,32
### Table 2. Summary of Intervention Studies on Treatment vs No Treatment for Gestational Diabetes (Key Questions 6 and 7)

| Source, country | Study design | Quality | Age, y | BMI | Ethnic majority | Inclusion criteria | Timing of randomization | Intervention components | Insulin/medication requirements | Gestational age at birth, mean (SD), wk |
|----------------|--------------|---------|--------|-----|----------------|-------------------|----------------------|-------------------------|-------------------------------|----------------------------------|
| Bevier et al, 1999 | RCT (n = 103; 83 analyzed [35 vs 48]) | Fair (no blinding and 19.5% incomplete outcome data) | Group 1: 26.3 (6.0) | Group 1 weight: 68.2 (11.4) kg | Group 2 weight: 72.4 (12.0) kg | Group 1 (HbA<sub>1c</sub> at 28 wGA): 4.7% (0.6%) | 24-28 wGA | Group 1: Diet, SMBG, and insulin if needed; random blood glucose measurement weekly and HbA<sub>1c</sub> testing at 28 and 32 wk | Group 1: 1 of 35 | Group 1: 39.6 (1.3) Group 2: 39.4 (1.5) |
| Bonomo et al, 2005 | RCT (n = 300; 300 analyzed [150 vs 150]) | Fair (no blinding) | Group 1: 31.1 (4.7) | Group 1: 23.1 (4.4) | Group 2: 30.7 (5.1) | OGCT: 8.44 (0.89) mmol/L | OGCT-positive and OGTT-negative on Carpenter and Coustan | Group 1: Diet, SMBG, biweekly blood work including FPG and HbA<sub>1c</sub> | Group 1: 39.4 (1.2) Group 2: 39.6 (1.7) |
| Crowther et al, 2005 | RCT (n = 1000; 1000 analyzed [490 vs 510; 506 vs 524 infants]) | Good | Group 1: 30.9 (5.4) | Group 1: 26.8 (KQ, 23.3-31.2) | Group 2: 30.1 (5.5) | Group 2: median, 26.0 (KQ, 22.9-30.9) | FPG: 86 (12.6) 2-h: 153.2 (14.4) | 24-34 wGA | Group 1: Diet, SMBG 4 times daily, insulin as needed | Group 1: 39.0 (IQR, 38.1-40.4); P = .01 |
| Deveer et al, 2013 | Nonrandomized controlled intervention study (n=100; 100 analyzed [50 vs 50]) | Fair (no blinding or allocation concealment; inadequate sequence generation) | Group 1: 29.5 (5.8) | Group 1: 28.0 (3.6) | Group 2: 29.1 (4.8) | Group 1: 153.2 (28.8) | OGCT-positive and OGTT-negative | Group 1: Diet | Group 1: 38.7 (1.2) Group 2: 38.9 (1.1) |

(continued)
| Source, country | Study design | Quality | Age, y | BMI* | Glycemic status at enrollment, mean (SD) mg/dL or % | Ethnic majority | Inclusion criteria (level of glycemia and others as relevant) | Timing of randomization | Intervention components | Insulin/medication requirements | Gestational age at birth, mean (SD), wk |
|----------------|-------------|---------|--------|------|-----------------------------------------------|---------------|-------------------------------------------------|------------------------|--------------------------|-------------------------------|----------------------------------|
| Fadl et al,28 2015 | RCT (n = 72; 69 analyzed [33 vs 36; 34 with exclusion of early miscarriage]) | Good (fair for outcomes with potential selective outcome reporting [shoulder dystocia, neonatal hypoglycemia, preterm deliveries]) | 1: 32.6 (5.9) | Group 1: 31.3 (6.4) | 71% Nordic | OGTT before 34 wGA (criteria 1 + risk factor or RBG >9.0 mmol/L); 75-g capillary OGTT: fasting <126 mg/dL or 2-h value ≥180 to <220 mg/dL | If early RBG >9 mmol/L, then given early OGTT (n = NR), if normal RBG value, then OGTT at 28-32 wGA | Group 1: Diet, SMBG 4 times daily, insulin as needed | Insulin: Group 1: 67% Group 2: NR | Group 1: 275 d (range, 258-288) Group 2: 273 d (range, 221-209) |
| Kokanali et al,31 2014 | RCT (n = 201; 201 analyzed [99 vs 102]) | Fair (blinding and allocation concealment NR) | At delivery: Group 1: 27.9 (5.8) | Group 2: 27.8 (5.8) | NR | OGTT-positive and 1 abnormal value by Carpenter and Coustan criteria | 24-28 wGA | Group 1: Diet therapy with dietician, SMBG (details NR), insulin as needed | Insulin: Group 1: NR Group 2: NR | Group 1: 269.1 (12.5) d Group 2: 286.8 (13.4) d |
| Landon et al,32 2009 | RCT (n = 958; 931 for most outcomes except hypoglycemia [n = 738; 77%]) | Good | 1: 29.2 (5.7) | Group 1: 30.1 (5.0) | 57% Hispanic | Between 24-31 wGA; >135 on OGCT; FPG <95 mg/dL and 2 or 3 abnormal on Carpenter and Coustan OGTT Excluded women with chronic hypertension, previous gestational diabetes, stillbirth | 24-31 wGA; mean, 28.8 (SD, 1.6) | Group 1: Diet, SMBG, insulin as needed (250% of fasting or postprandial levels elevated) Group 2: Routine care, random glucose measurement at primary care provider discretion | Insulin: Group 1: 7.6% Group 2: 0.4% | Group 1: 39.0 (1.8) Group 2: 38.9 (1.8) |
| Garner et al,29 1997 | RCT (n = 300; 299 analyzed [149 vs 150]) | Fair (insufficient blinding of patients) | Prepregnancy weight: Group 1: 68.9 (16.9) kg Group 2: 71.2 (19.8) kg | 75-g OGCT: 182.0 (25.6) | 91% White | Positive 75-g OGCT and gestational diabetes criteria (FPG, 4.8 mmol/L; 1-h, 10.9 mmol/L; 2-h, 9.6 mmol/L [number abnormal, NR]) diagnosed between 24-32 wGA; otherwise, low-risk pregnancy Excluded women with chronic hypertension | 24-32 wGA | Group 1: Tertiary care center follow-up with obstetrician and endocrinologist; diet, daily SMBG, biweekly fetal monitoring, insulin as needed (13 [7.8%] met type 2 diabetes criteria) Group 2: Primary care provider; twice weekly SMBG (results sent to independent observer); no fetal monitoring unless indicated (16 [10.6%] meeting type 2 diabetes criteria were given treatment) | Insulin: Group 1: NR | Group 2: NR but 10.6% type 2 diabetes |

(continued)
Table 2. Summary of Intervention Studies on Treatment vs No Treatment for Gestational Diabetes (Key Questions 6 and 7) (continued)

| Source, country | Study design | Quality | Age, y | BMI* | Glycemic status at enrollment, mean (SD) mg/dL or % | Inclusion criteria (level of glycemia and others as relevant) | Ethnic majority | Timing of randomization | Intervention components | Insulin/medication requirements | Gestational age at birth, mean (SD), wk |
|-----------------|--------------|---------|--------|------|-----------------------------------------------|-------------------------------------------------------------|----------------|------------------------|------------------------|-----------------------------|-------------------------------|
| Yang et al,15 2014 China | RCT (n = 948 [130 vs 112 excluded from break in protocol from site renovations]; 700 analyzed [361 vs 339]) | Fair (unclear sequence generation; no blinding of patients or primary care providers) | Group 1: 29.9 (3.5) | Group 2: 29.7 (3.2) | Prepregnancy: Group 1: 22.9 (3.6) Group 2: 23.4 (3.9) OGTT: Group 1: fasting, 91.9 (10.8); 1-h, 182.0 (25.2); 2-h, 151.4 (21.6) Group 2: fasting, 90.1 (9.0); 1-h, 180.2 (23.4); 2-h, 151.4 (25.2) | Gestational diabetes diagnosed with 2-step IADPSG 2010 criteria (with 50-g OGCT) (not meeting type 2 diabetes criteria using FPG and HbA1c) Excluded women with chronic hypertension | 97% Han Chinese | 24-29 wGA; mean, 26.3 (SD, 1.4) | Group 1: Shared care system (primary care hospital then obstetric hospital) with team of nurses and physicians; diet, physical activity, SMBG Group 2: One hospital-based education session by diabetes educator (diet and physical activity but no SMBG); insulin if HbA1c >6.5% at 34 wk | | |
| Hughes et al,30 2018 New Zealand | RCT (n = 47; 44 analyzed [23 vs 21]) | Fair (unclear for baseline imbalances; no blinding) | Age at expected delivery date: Group 1: median, 30.5 (IQR, 28.0-34.5) Group 2: median, 32.0 (29.5-36.0) | Baseline: Group 1: median, 29.6 (IQR, 24.1-35.6) Group 2: median, 30.3 (IQR, 27.1-38.4) HbA1c at booking: Group 1: median, 42 (Q1, 41-45) mmol/mol; mean, 6.0% (SD, 2.4%) Group 2: median, 42 (IQR, 41-45) mmol/mol; mean, 6.0% (SD, 2.4%) | 51% Asian | HbA1c, 5.9%-6.4% (41-46 mmol/mol) at booking | <14 wGA | Group 1: Diabetes clinic and lead maternity carer (midwife or obstetrician): ongoing lifestyle education, daily SMBG (before and after each meal), and medication as required (metformin and/or insulin) Group 2: Standard care with lead maternity caregiver and 75-g OGTT screening at 24 wGA (IADPSG or New Zealand criteria: FBG ≥5.5 mmol/L [99 mg/dL]) or 2-h blood glucose ≥9.0 mmol/L [162 mg/dL]), with referral if gestational diabetes | Group 1: 17/23 (metformin in 14 and insulin in 15 women) (all before 24 wk) Group 2: 11/22 (metformin in 3 and insulin in 11 women) | | |

Early treatment vs usual care

| Hughes et al,30 2018 New Zealand | RCT (n = 47; 44 analyzed [23 vs 21]) | Fair (unclear for baseline imbalances; no blinding) | Age at expected delivery date: Group 1: median, 30.5 (IQR, 28.0-34.5) Group 2: median, 32.0 (29.5-36.0) | Baseline: Group 1: median, 29.6 (IQR, 24.1-35.6) Group 2: median, 30.3 (IQR, 27.1-38.4) HbA1c at booking: Group 1: median, 42 (Q1, 41-45) mmol/mol; mean, 6.0% (SD, 2.4%) Group 2: median, 42 (IQR, 41-45) mmol/mol; mean, 6.0% (SD, 2.4%) | 51% Asian | HbA1c, 5.9%-6.4% (41-46 mmol/mol) at booking | <14 wGA | Group 1: Diabetes clinic and lead maternity carer (midwife or obstetrician): ongoing lifestyle education, daily SMBG (before and after each meal), and medication as required (metformin and/or insulin) Group 2: Standard care with lead maternity caregiver and 75-g OGTT screening at 24 wGA (IADPSG or New Zealand criteria: FBG ≥5.5 mmol/L [99 mg/dL]) or 2-h blood glucose ≥9.0 mmol/L [162 mg/dL]), with referral if gestational diabetes | Group 1: 17/23 (metformin in 14 and insulin in 15 women) (all before 24 wk) Group 2: 11/22 (metformin in 3 and insulin in 11 women) | | |

(continued)
Table 2. Summary of Intervention Studies on Treatment vs No Treatment for Gestational Diabetes (Key Questions 6 and 7) (continued)

| Source, country | Study design | Quality | Age, y | BMI* | Glycemic status at enrollment, mean (SD) mg/dL or % | Ethnic majority | Inclusion criteria (level of glycemia and others as relevant) | Timing of randomization | Intervention components | Insulin/medication requirements | Gestational age at birth, mean (SD), wk |
|-----------------|--------------|---------|--------|------|--------------------------------------------------|----------------|-----------------------------------------------------------------|------------------------|--------------------------|----------------------------------|----------------------------------|
| Osmundsen et al,33 2016 US | RCT (n = 95 [50 vs 45]; 83 analyzed [42 vs 41]; 74 with delivery data [37 vs 37]) | Fair (no blinding, incomplete outcome data, and some possible selective outcome reporting) | Group 1: 32.4 (5.1) | Group 2: 34.3 (5.2) | Prepregnancy: Group 1: median, 27.2 (IQR, 24.8-33.2) Group 2: median, 27.4 (IQR, 22.6-32.7) | 45% Hispanic; 37% Asian | HbA1c: 5.7%-6.4% before 14 wGA Excluded women with a prior infant with birth injury or shoulder dystocia possibly attributable to diabetes, or prior macrosomic infant | <14 wGA (mean, 11.1 wk) | Group 1: Diet with certified diabetes educator; SMBG 4 times daily, insulin as needed; OGTT at 26-28 wk, with women with negative results continuing dietary but reduced SMBG | Group 1: 135.9% Group 2: 26.3% | Group 1: 38.3 (2.3) Group 2: 38.2 (2.0) |
| Simmons et al,34 2018 New Zealand | RCT (n = 21; 20 analyzed [11 vs 9]) | Good | Group 1: 29 (5) | Group 2: 30 (7) | Early (<20 wGA) OGTT results: Group 1: fasting, 91.9 (7.2) mmol/L; 1-h, 144.1 (30.6) mmol/L; 2-h, 126.1 (34.2) mmol/L Group 2: fasting, 93.7 (5.4) mmol/L; 1-h, 151.4 (28.8) mmol/L; 2-h, 122.5 (30.6) mmol/L | 55% White | With risk factors and gestational diabetes on 75 g OGTT by IADPSG criteria, <20 wGA | 4-20 wGA | Group 1: Education, diet, SMBG, metformin or insulin as needed Group 2: Routine prenatal care, with screening at 24-28 wGA | Insulin or metformin: Group 1: 36% Group 2: 40% | Group 1: 38.7 (1.4) Group 2: 39.2 (0.6) |

(continued)
Table 2. Summary of Intervention Studies on Treatment vs No Treatment for Gestational Diabetes (Key Questions 6 and 7) (continued)

| Source, country | Study design | Quality | Age, y | BMI* | Glycemic status at enrollment, mean (SD) mg/dL or % | Inclusion criteria (level of glycemia and others as relevant) | Timing of randomization | Intervention components | Insulin/medication requirements | Gestational age at birth, mean (SD), wk |
|-----------------|--------------|---------|--------|------|---------------------------------------------------|-----------------------------------------------------------|-------------------------|--------------------------|-------------------------------|----------------------------------------|
| Vinter et al,²⁷ 2018 Denmark | Nonrandomized controlled intervention study of gestational diabetes (n = 90; 90 analyzed [36 vs 54]) | Fair (not randomized) | 29.0 (4.4) | Prepregnancy or first trimester: 34.5 (4.3) | Venous fasting: 93.7 (3.6) Capillary 2-h: 117.1 (19.8) (first trimester) | BMI 30-40 (prepregnancy or first measured weight in pregnancy); diagnosed retrospectively with gestational diabetes by modified WHO 2013 criteria in early pregnancy (12-15 wGA); venous FPG ≥5.1 mmol/L and/or 2-h capillary ≥8.5 mmol/L, but not meeting Danish criteria for gestational diabetes (2-h capillary ≥9.0 mmol/L) at any time (12-15, 28-30, or 34-36 wGA) | 12-15 wGA | Group 1: Lifestyle intervention: 4 diet counseling sessions with a trained dietician, encouraged to perform 30-60 min daily exercise with a free full membership to a fitness center for 6 mo until delivery (included closed exercise classes with a physiotherapist 1 h weekly); no SMBG or insulin assessment per protocol Group 2: Routine care Both groups were monitored with fasting blood samples, OGTTs, sonographic fetal biometry, and measurements of maternal weight and blood pressure | Group 1: NR Group 2: NR | Group 1: 40 (range, 39-41.3) Group 2: 40.7 (range, 39-41.3) |

Abbreviations: BMI, body mass index; FBG, fasting blood glucose; FPG, fasting plasma glucose; HbA₁c, glycated hemoglobin; IADPSG, International Association of Diabetes and Pregnancy Study Groups; IQR, interquartile range; NR, not reported; OGCT, oral glucose tolerance test; OGTT, oral glucose challenge test; RBG, random blood glucose;RCT, randomized clinical trial; SGA, small for gestational age; SMBG, self-monitoring blood glucose; wGA, weeks of gestation; WHO, World Health Organization.

SI conversion factor: To convert glucose values from mmol/L to mg/dL, divide by 0.0555.

* Calculated as weight in kilograms divided by height in meters squared.
neonatal hypoglycemia (total [5 trials], or requiring intravenous treatment [2 trials], hyperbilirubinemia (5 trials), or APGAR scores (2 trials), though results were often heterogeneous, imprecise, or both.

Long-term follow-up of 1 trial found no significant association between treatment for gestational diabetes vs no treatment and maternal impaired fasting glucose, obesity, metabolic syndrome, or type 2 diabetes at 5 to 10 years. No study measured effects of treatment on long-term quality of life, cardiovascular outcomes, or mortality or major morbidity from type 2 diabetes.

Regarding long-term child outcomes, treatment of mothers for gestational diabetes was not significantly associated with reduced risk of overweight/obesity at 4 to 7 years (2 trials), obesity at 7 to 9 years (2 trials), impaired glucose tolerance (median, 9 years [1 trial] or impaired fasting glucose (median, 7-9 years [2 trials]). Evidence from 2 RCTs on long-term risk of type 2 diabetes in children was too sparse to determine effect of treatment.

Subgroup analyses from 1 trial found no significant differences in effects of gestational diabetes treatment for several maternal and fetal outcomes based on timing of treatment initiation, race/ethnicity, severity of dysglycemia, or BMI. Differences in gestational diabetes diagnostic criteria did not appear to affect findings or explain inconsistency.

Early Treatment vs Usual Care
Findings from 4 small trials (n = 21-95) of treatment for gestational diabetes in early pregnancy (using HbA1c concentration or IADPSG criteria before 14 to 15 weeks of gestation) were highly imprecise.
Risk-based screening (75-g 2-h OGTT NICE criteria) was associated with a reduced risk of late (≥28 weeks’ gestation) stillbirth (adjusted OR, 0.68 [95% CI, 0.47-0.97])

Universal 2-step screening (50-g OGCT and 75-g 2-h OGTT using IADPSG), with women having risk factors screened in first trimester (51% of screened), associated with a reduced risk of cesarean deliveries (ARR, 5%), birth injuries (1%), and admissions to the NICU (8%) compared to women without risk factors.

For NICU admissions, effects for women screened in first trimester were larger than for those screened later.

Two small studies from the prior review focused on selected populations and showed no associations with screening.

Evidence from observational studies on harms of screening (2 studies) or a gestational diabetes diagnosis (5 studies) was limited but suggested that undergoing screening or receiving a false-positive result may not be associated with anxiety or depression; receiving a gestational diabetes diagnosis may result in a small, transient increase in anxiety symptoms; and that the diagnosis may have some adverse labeling effects affecting delivery management and hospital experiences associated with breastfeeding.

Harms of screening: reasonably consistent; some imprecision

Harms of gestational diabetes diagnosis: reasonably consistent (labeling); unknown consistency (anxiety)

Observational studies; not intention/offer-to-screen designs

Findings on hospital experiences may be confounded by hospital policies, gestational diabetes treatment, and intentions before delivery

Low for no association between undergoing screening and anxiety and depression symptoms

Studies from Canada and Australia with predominately White women; screening used the OGCT

**Table 3. Summary of Evidence**

| Comparison | Studies; observations (No.); study designs | Summary of findings | Consistency and precision | Other limitations | Strength of evidence | Applicability |
|------------|------------------------------------------|---------------------|--------------------------|-------------------|---------------------|--------------|
| **KQ1: Benefits of screening** | | | | | | |
| Screening vs no screening | | | | | | |
| Prior report: 2 retrospective cohort studies (n = 544) Update: 1 case-control and 1 retrospective cohort (n = 3792) | Risk-based screening (75-g 2-h OGTT NICE criteria) was associated with a reduced risk of late (≥28 weeks’ gestation) stillbirth (adjusted OR, 0.68 [95% CI, 0.47-0.97]) | Consistency unknown, with 1 study for each outcome | Reasonably precise for stillbirth, cesarean deliveries, birth injuries, and NICU admissions; some imprecision for macrosomia, hypoglycemia, or hyperbilirubinemia | Observational studies without intention/offer to screen designs Some concerns about selection biases and confounding Selective outcome or analysis reporting not detected | Insufficient | Findings mainly applicable to screening approaches with targeted screening for women with risk factors |
| | Universal 2-step screening (50-g OGCT and 75-g 2-h OGTT using IADPSG), with women having risk factors screened in first trimester (51% of screened), associated with a reduced risk of cesarean deliveries (ARR, 5%), birth injuries (1%), and admissions to the NICU (8%) compared to women without risk factors. | | | | | |
| | For NICU admissions, effects for women screened in first trimester were larger than for those screened later. | | | | | |
| | Two small studies from the prior review focused on selected populations and showed no associations with screening. | | | | | |
| **KQ2: Harms of screening** | | | | | | |
| Screening vs no screening and gestational diabetes vs no gestational diabetes | | | | | | |
| Prior report: 0 Update: 5 cohort studies and 2 cross-sectional studies (n = 166082) | Evidence from observational studies on harms of screening (2 studies) or a gestational diabetes diagnosis (5 studies) was limited but suggested that undergoing screening or receiving a false-positive result may not be associated with anxiety or depression; receiving a gestational diabetes diagnosis may result in a small, transient increase in anxiety symptoms; and that the diagnosis may have some adverse labeling effects affecting delivery management and hospital experiences associated with breastfeeding. | | | | | |
| | Harms of screening: reasonably consistent; some imprecision |
| | Harms of gestational diabetes diagnosis: reasonably consistent (labeling); unknown consistency (anxiety) |
| | Observational studies; not intention/offer-to-screen designs |
| | Findings on hospital experiences may be confounded by hospital policies, gestational diabetes treatment, and intentions before delivery |
| | Low for no association between undergoing screening and anxiety and depression symptoms |
| | Studies from Canada and Australia with predominately White women; screening used the OGCT |
| **KQ3: Comparative effectiveness of screening strategies** | | | | | | |
| One-step IADPSG vs 2-step Carpenter and Coustan screening | | | | | | |
| Prior report: 0 Update: 5 RCTs (n = 25772) | One large RCT (n = 23792) accounted for 92% of patients | Pregnancy outcomes: consistent and precise for hypertensive disorders, total cesarean deliveries, and induction of labor | Large RCT had substantial crossover (>25% of IADPSG group received Carpenter and Coustan for diagnosis), but findings were very similar in analysis accounting for adherence Possible selective outcome or analysis reporting not detected | Large RCT could not be explained despite seeking author contact | Insufficient | Four trials conducted in US, with fairly diverse populations Comparison highly applicable to US |
| | Pregnancy outcomes: no significant association with primary cesarean deliveries, preeclampsia, hypertensive disorders, gestational hypertension, total cesarean deliveries, induction of labor, preterm birth, and maternal birth trauma |
| | Fetal/neonatal outcomes: No significant association for mortality, birth injury, shoulder dystocia, macrosomia, neonatal hypoglycemia, neonatal hyperbilirubinemia, NICU admissions, neonatal respiratory distress, Apgar scores at <7 min, or SGA |
| | The large trial reported a significant increase neonatal hypoglycemia from 1-step screening |
| | Long-term outcomes: No data |
| | Pregnancy outcomes: consistent and precise for hypertensive disorders, total cesarean deliveries, and induction of labor |
| | Large RCT had substantial crossover (>25% of IADPSG group received Carpenter and Coustan for diagnosis), but findings were very similar in analysis accounting for adherence Possible selective outcome or analysis reporting not detected | Large RCT could not be explained despite seeking author contact |
| | Pregnancy outcomes: moderate for no significant association with total cesarean deliveries, induction of labor, primary cesarean deliveries, preterm birth and hypertensive disorders |
| | Insufficient for preeclampsia, gestational hypertension, and maternal birth trauma |
| | Fetal/neonatal outcomes: moderate for no significant association with mortality, birth injury, shoulder dystocia, macrosomia, hyperbilirubinemia, SGA, LGA, and NICU admissions |
| | Low for no significant association with neonatal hypoglycemia |
| | Insufficient for Apgar scores |
Table 3. Summary of Evidence (continued)

| Comparison | Study designs | Summary of findings | Consistency and precision | Other limitations | Strength of evidence | Applicability |
|------------|--------------|---------------------|---------------------------|------------------|----------------------|--------------|
| Early vs usual timing for Carpenter and Coustan screening | Prior report: 1 RCT (n = 922) | Pregnancy outcomes: Preeclampsia (RR 1.42 [95% CI, 0.99-2.05]; ARD, 4.0%); no significant association for gestational hypertension, hypertensive disorders in pregnancy, primary cesarean delivery, induction of labor | Pregnancy outcomes: Consistency unknown; some imprecision | No concerns | Pregnancy outcomes: Low for more preeclampsia (not significant) and no significant association for other outcomes | US trial with mostly Black and Hispanic population; 100% obese; excluded women with prior cesarean delivery; comparison highly applicable |
| | | Fetal/neonatal outcomes: No significant association for shoulder dystocia, macrosomia, LGA, hypoglycemia, hyperbilirubinemia | Consistency unknown; some imprecision | | | |
| | | Long-term outcomes: No data | Long-term outcomes: No data | | | |
| | | | | | | |
| KQ4: Diagnostic accuracy of screening tests | | Pooled estimates: 140 mg/dL: sensitivity, 81.9% (95% CI, 71.2%-91.1%); specificity, 81.8% (95% CI, 75.3%-88.6%); 135 mg/dL: sensitivity, 83.3% (95% CI, 72.2%-94.9%); specificity, 81.3% (95% CI, 74.3%-88.1%) | Consistent and precise | Half of the studies for each analysis were fair quality, but this did not appear to influence findings | Moderate (140 mg/dL) and low (135 mg/dL) for good accuracy; insufficient for 130 mg/dL | Studies varied widely in country of origin; screening and diagnostic test highly applicable |
| 50-g OGCT vs Carpenter and Coustan | Prior report: 5 studies (n = 5501) Update: 8 studies (n = 6190) | | Not pooled: 130 mg/dL: sensitivity, 75%-100%; specificities, 25%-86% | | | |
| 50-g OGCT vs NDDG | Prior report: 6 studies (n = 5375) Update: 0 | Pooled estimates: 140 mg/dL: sensitivity, 85.0% (95% CI, 72.0%-92.6%); specificity, 81.2% (95% CI, 75.3%-85.6%) | Consistent and precise | Four of 6 studies were good quality, and quality did not appear to influence findings | Moderate (140 mg/dL) and low (135 mg/dL) for good accuracy; insufficient for 130 mg/dL | See 50-g OGCT vs Carpenter and Coustan |
| 50-g OGCT vs IADPSG | Prior report: 0 Update: 2 studies (n = 2091) | Not pooled: Sensitivity: low (<70%) across all cutoffs | Consistent and precise | No concerns | Moderate for poor accuracy | See 50-g OGCT vs Carpenter and Coustan |

(continued)
| Comparison | Studies; observations (No.); study designs | Summary of findings | Consistency and precision | Other limitations | Strength of evidence | Applicability |
|------------|------------------------------------------|---------------------|--------------------------|------------------|---------------------|--------------|
| **Fasting plasma glucose vs Carpenter and Coustan** | Prior report: 4 studies (n = 6889) Update: 3 studies (n = 1972) | Pooled estimates: FPG 79 mg/dL: sensitivity, 96% (95% CI, 92%-98%); specificity, 35% (95% CI, 30%-41%) FPG 85 mg/dL: sensitivity, 88% (95% CI, 84%-91%); specificity, 73% (95% CI, 46%-90%) FPG 90 mg/dL: sensitivity, 81% (95% CI, 75%-85%); specificity, 82% (95% CI, 61%-93%) FPG 95.5 mg/dL: sensitivity, 58% (95% CI, 32%-81%); specificity, 98% (95% CI, 88%-100%) Not pooled: Across all cutoffs, sensitivity appeared fairly high (>90%) using cutoffs ≤80 mg/dL and specificity appeared high (≥90%) using cutoffs >90 mg/dL. | 79, 85, and 90 mg/dL: Sensitivity consistent and precise; some inconsistency for specificity ≤80 mg/dL: Consistent (but most thresholds only reported by 2 studies) and precise for high sensitivity | Two studies included in pooled estimates used selective populations (positive on OGCT or with clinical risk factors), which may have affected findings | Low (85 and 90 mg/dL) for good accuracy; low for high sensitivity (to rule out) with ≤80 mg/dL and specificity (to rule in) with >90 mg/dL | See 50-g OGCT vs Carpenter and Coustan |
| **Fasting plasma glucose vs IADPSG** | Prior report: 0 Update: 9 studies (n = 59 278) | Pooled estimate at 24 weeks of gestation or greater: FPG 90 mg/dL: sensitivity, 79% (95% CI, 65%-89%); specificity, 96% (95% CI, 95%-97%) | At 24 weeks or greater: FPG 90 mg/dL: Some inconsistency but precise for sensitivity FPG ≤80 mg/dL: Consistent (but most thresholds only reported by 2 studies) and precise for high sensitivity Early screening: Inconsistent sensitivity | 6 of 9 studies were fair quality, but quality did not appear to influence findings | Moderate (90 mg/dL at 24 weeks) for good accuracy; low (≤80 mg/dL) to rule out gestational diabetes; low for low accuracy when screening before 24 weeks | Studies varied in country, and findings appear to be applicable to a diverse population; 90 mg/dL is very similar to the diagnostic value for FPG in this criteria, which requires only 1 abnormal value |
| **HbA1c** | Prior report: 3 studies (n = 1075) Update: 15 studies (n = 9413) | Against each criteria and for each time point, 1 or 2 studies contributed data for most thresholds At no threshold were sensitivity and specificity both high enough for use as a primary screening test Sensitivity >90% at cutoffs of 4.5% to 5.0% (Carpenter and Coustan and NDDG) and ≤4.7% (IADPSG) in second trimester and 4.5% to 4.8% (NDDG) in first trimester; may allow ruling out | Some inconsistency and imprecision | Most studies limited due to poor reporting on patient selection, selection of cutoffs, and fasting protocols | Low for poor accuracy across thresholds; low for ≤5.5% (Carpenter and Coustan and NDDG) and ≤4.7% (IADPSG) in second trimester and 4.5% to 4.8% (NDDG) in first trimester to rule out gestational diabetes | See 50-g OGCT vs Carpenter and Coustan |
| **Risk-based screening** | Prior report: 2 studies (n = 1912) Update: 1 study (n = 258) | Three studies compared different models with Carpenter and Coustan, NDDG, and IADPSG criteria; for Carpenter and Coustan and IADPSG studies incorporated FPG, which seemed to increase sensitivity All screening still used either FPG or OGCT Sensitivity may be high enough (82%-98%) to rule out gestational diabetes; specificity (17%-80%) too low to replace OGCT | Single studies for each tool and criteria; some imprecision | No concerns; all studies used validation cohorts | Low for poor accuracy for primary screening test; may allow ruling out | Studies from Brazil, Canada, and Austria; unknown how many clinicians use risk-based screening |
### Table 3. Summary of Evidence (continued)

| Comparison | Summary of findings | Consistency and precision | Other limitations | Strength of evidence | Applicability |
|------------|---------------------|---------------------------|------------------|----------------------|---------------|
| **KQ6: Benefits of treatment** | | | | | |
| Treatment for gestational diabetes at 24 weeks of gestation vs no treatment | | | | | |
| Prior report: 5 RCTs; Update: 3 RCTs and 1 nonrandomized intervention study (n = 3982) | Pregnancy outcomes: Primary cesarean delivery: RR, 0.70 (95% CI, 0.54-0.91); ARD, 5.3%
Preterm delivery: RR, 0.75 (95% CI, 0.56-1.01); ARD, 2.3%
No significant association with hypertensive disorders of pregnancy, gestational hypertension, total cesarean delivery, induction of labor, maternal birth trauma
Fetal/neonatal outcomes:
- Birth injury: Peto OR, 0.33 (95% CI, 0.11-0.99); ARD, 0.2%
- Shoulder dystocia: RR, 0.42 (95% CI, 0.23-0.77); ARD, 1.3%
- Macrosomia >4000 g: RR, 0.53 (95% CI, 0.41-0.68); ARD, 8.9%
- LGA: RR, 0.56 (95% CI, 0.47-0.66); ARD, 8.4%
- NICU admissions: RR, 0.73 (95% CI, 0.53-0.99); ARD, 2.0%
- No significant association with mortality, macrosomia >4500 g, respiratory distress syndrome, any hypoglycemia, hyperbilirubinemia, APGAR scores
Long-term outcomes: No significant differences in childhood overweight (BMI ≥85th percentile), obesity (BMI ≥95th percentile), metabolic impairment, type 2 diabetes, or maternal obesity (BMI ≥30) or metabolic impairment (impaired fasting glucose), the metabolic syndrome (5-10 y), or type 2 diabetes (5-10 y)
Subgroups: No significant interactions based on timing of treatment initiation, criteria for diagnosis/glycemic severity, BMI (only assessed for LGA), or race and ethnicity
Sensitivity analyses removing 3 trials with eligibility based on screening positive but no gestational diabetes did not affect conclusions; one new trial enrolled women with gestational diabetes based on IADPSG criteria but FPG was higher and 2-h postload glucose levels similar to other trials in the prior review, so this did not explain any inconsistency in effect | Consistent and precise for macrosomia >4000 g and LGA
Inconsistent and imprecise for preeclampsia, birth injury, and mortality
Inconsistent for gestational hypertension, primary cesarean delivery, emergency cesarean delivery, shoulder dystocia
Some inconsistency for induction of labor and shoulder dystocia
Large inconsistency for hypertension disorders
Unknown consistency and large imprecision for childhood and maternal metabolic impairment and development of type 2 diabetes | Some concern for total cesarean delivery, induction of labor, and NICU admissions from open-label designs
Studies of long-term outcomes had high rates of attrition | High for reduced risk of macrosomia >4000 g and LGA
Moderate for reduced risk of primary cesarean delivery, shoulder dystocia, and NICU admissions and for no association with gestational hypertension, total cesarean deliveries, maternal birth trauma, respiratory distress syndrome, hypoglycemia, hyperbilirubinemia
Low for reduced risk of preeclampsia, preterm labor, birth injury and for no association with hypertensive disorders, emergency cesarean delivery, induction of labor, mortality, macrosomia >4500 g, and childhood obesity
Insufficient for childhood and maternal metabolic impairment and development of type 2 diabetes | Trials from various countries; 2 from the US enrolled 97% and 57% Hispanic women with findings similar to the conclusions
Most data from 3 large trials with 2-step screening for gestational diabetes diagnosis
Eligibility criteria included singleton pregnancies for 12 trials, women without chronic hypertension in 4 trials, and women without previous gestational diabetes in the largest 2 trials |
| Comparison | Studies; observations (No.; study designs) | Summary of findings | Consistency and precision | Other limitations | Strength of evidence | Applicability |
|------------|------------------------------------------|---------------------|--------------------------|------------------|----------------------|---------------|
| Early gestational diabetes treatment vs usual care | Prior report: 0 Update: 3 RCTs and 1 nonrandomized intervention study (n = 253) | Pregnancy outcomes: No significant associations for preeclampsia, gestational hypertension, hypertensive disorders of pregnancy, cesarean delivery, primary cesarean delivery, emergency cesarean delivery, induction of labor, preterm delivery, excessive gestational weight gain Fetal/neonatal outcomes: No significant associations for mortality, birth injury, shoulder dystocia, macrosomia >4000 g, macrosomia >4500 g, LGA, NICU admissions, any hypoglycemia, hyperbilirubinemia Long-term outcomes: No data Subgroups: Interactions between BMI and early treatment vs usual care imprecise | Highly imprecise for all outcomes | Some concern for total cesarean delivery, induction of labor, and NICU admissions from open-label designs Studies of long-term outcomes had high rates of attrition | Insufficient for all outcomes of early treatment | Trials from Australia, New Zealand, Denmark and the US, largely nonminority populations |

KQ7: Harms of treatment

| Treatment for gestational diabetes at 24 weeks’ gestation or later vs no treatment | Prior report: 5 trials Update: 4 trials (n = 3982) | Pregnancy outcomes: No significant association with severe maternal hypoglycemia Large association with reduced risk of macrosomia (>4000 g; RR, 0.53 [95% CI, 0.41-0.68]) but no significant association with risk of total cesarean deliveries (RR, 0.95 [95% CI, 0.83-1.08]); cesarean delivery may be associated with gestational diabetes Fetal/neonatal outcomes: No significant association with SGA, low birth weight, neonatal hypoglycemia requiring IV glucose therapy Long-term outcomes: No data Subgroups: No effect of SGA based on ethnicity or glycemic status | Highly imprecise for maternal hypoglycemia Some imprecision and inconsistency for severe neonatal hypoglycemia (requiring IV treatment) Some imprecision for SGA | No concerns; results were consistent with those from 2 large good quality trials | Moderate for no association with SGA Low for no association with severe neonatal hypoglycemia Insufficient for severe maternal hypoglycemia | See KQ6 |

Early gestational diabetes treatment vs usual care | Prior report: 0 Update: 3 trials (n = 123) | No significant association with SGA | Highly imprecise | Open-label in 3 trials; 1 was not randomized and 1 had high attrition | Insufficient | See KQ6 |

Abbreviations: ARD, absolute risk difference; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); FPG, fasting plasma glucose; HbA_1c, glycated hemoglobin; IADPSG, International Association of Diabetes and Pregnancy Study Groups; KQ, key question; LGA, large for gestational age; NDDG, National Diabetes Data Group; NGT, normal glucose tolerance; NICU, neonatal intensive care unit; OGCT, oral glucose challenge test; OGTT, oral glucose tolerance test; OR, odds ratio; RCT, randomized clinical trial; RR, relative risk; SGA, small for gestational age. SI conversion factor: To convert glucose values to mmol/L, multiply by 0.0555.
Key Question 7. What are the harms of treatment of gestational diabetes, including severe maternal and neonatal hypoglycemia, delivery of neonates who are small for gestational age, and poor long-term growth and development outcomes in the child?

Treatment offered at 24 weeks of gestation or later was not significantly associated with increased risk of SGA (6 trials; n = 2646; RR, 1.10 [95% CI, 0.83 to 1.47]). The overall risk in this review was lower than in previous reviews.25,27,31,32,36 No trial reported on the association between treatment and poor long-term growth and development outcomes in childhood. Findings from small RCTs of early treatment vs usual care were imprecise or did not report harms (eg, maternal hypoglycemia).

Discussion

The findings in this evidence report are summarized in Table 3. Direct evidence on the benefits of screening vs no screening remains limited and consists of observational studies with methodological limitations. Few studies reported on harms from screening or a diagnosis of gestational diabetes and those available were limited by imprecision and methodological limitations. There were no significant associations between screening using 1-step IADPSG vs 2-step Carpenter and Coustan criteria, but some statistical heterogeneity was present (especially for neonatal hypoglycemia) and estimates were heavily weighted by 1 large trial20 that accounted for 92% of patients.

Treatment vs no treatment was associated with reduced risk for some pregnancy and several neonatal/fetal outcomes. Findings are most applicable for hyperglycemia identified using 2-step screening approaches and to adult (vs adolescent) women with singleton pregnancies and without chronic hypertension or previous gestational diabetes. Most of the treatment interventions relied on frequent self-monitoring of blood glucose levels and clinic visits to monitor glucose targets, which could reduce applicability of findings to women with limited or no insurance coverage, health care access, or ability to perform self-monitoring. Results for cesarean delivery and labor induction are difficult to interpret because of differences in delivery practices. Findings are sparse for long-term health outcomes from treatment and for all outcomes from early treatment. No trial of treatment at 24 weeks of gestation or after used oral medications; therefore, potential medication harms would not have been captured.

This review differs from the 2012 USPSTF review11 by including additional evidence on potential harms of screening and gestational diabetes diagnosis; evaluating comparative effectiveness of different screening strategies; and relying on more rigorous inclusion criteria and applicable comparisons for test accuracy. Although findings were generally consistent with those from the prior review, there are some differences. New evidence resulted in increased certainty regarding the accuracy of fasting plasma glucose and HbA1c levels as screening tests and the association between treatment and improved outcomes, including reduced risk of NICU admissions. Additional information on preeclampsia and NICU admissions was obtained from authors of 1 trial, enhancing handling of these data. Several publications from one of the larger treatment trials34 provided new evidence regarding lack of effect for several subgroups and long-term outcomes. For the new KQ on comparative effectiveness, several trials were located including 3 large trials19,20,24 and 1 very large trial20 from the US examining highly applicable comparisons. The greater prevalence in gestational diabetes diagnosis resulting from 1-step IADPSG vs 2-step Carpenter and Coustan screening, without associated benefits, suggests potential overdiagnosis and overtreatment. In addition, the 1-step approach requires additional resources related to having all women undertake a 2-hour OGTT and provision of counseling and treatment to more women.

Evaluating the effectiveness of screening vs no screening remains heavily reliant on indirect evidence about test accuracy and treatment effects. Although evidence on diagnostic accuracy is useful for assessing which screening tests may be most useful in a 2-step approach, reliance on these tests alone would result in a high number false-positive results (especially using lower cutoffs with high sensitivity), particularly in general-prevalence populations. The results of trials of treatment for women with positive OGCT screening results but not gestational diabetes,103 and for those with gestational diabetes by IADPSG criteria but excluding those with 2 abnormal glucose values,104 would be useful to further inform assessment of treatment benefits among women with lesser degrees of dysglycemia.

Limitations

This review had several limitations. First, only English-language studies were included.105 Second, graphical and statistical tests for small-sample effects were not conducted because all analyses included fewer than 10 trials.106

Third, the DerSimonian and Laird random-effects model was used to pool studies, which may result in CIs that are too narrow, particularly when heterogeneity is present.107 However, results were similar when analyses were repeated using the profile likelihood method. Fourth, the observational studies included for KQs for which trials were lacking were susceptible to unmeasured confounding and other methodological limitations.

Fifth, some studies were conducted in countries in which screening and treatment for gestational diabetes, as well as management of pregnancy, may differ from that in the US. However, this review focused on screening and diagnostic criteria used in the US, and results appeared consistent across geographic settings. Sixth, data on how the effects of screening and treatment varied according to patient characteristics such as race/ethnicity, age, and other socioeconomic factors were very limited. Seventh, studies that applied older definitions for gestational diabetes or that did not screen for preexisting diabetes2 may have included some women with overt diabetes, who are expected to have worse outcomes.108

Conclusions

Direct evidence on screening vs no screening remains limited. One vs 2-step screening was not significantly associated with improved health outcomes. At or after 24 weeks of gestation, treatment of gestational diabetes was significantly associated with improved health outcomes.
Accepted for Publication: June 15, 2021.

Author Contributions: Ms Pillay had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Pillay, Donovan, Chou, Hartling.

Acquisition, analysis, or interpretation of data: Pillay, Guirard, Zakher, M. Gates, A. Gates, Vandermeer, Bougatsos, Chou, Hartling.

Drafting of the manuscript: Pillay, Guirard, Zakher, M. Gates, A. Gates, Bougatsos, Chou, Hartling.

Critical revision of the manuscript for important intellectual content: Donovan, Vandermeer, Chou.

Statistical analysis: Pillay, Guirard, Zakher, M. Gates, A. Gates, Vandermeer.

Obtained funding: Chou, Hartling.

Administrative, technical, or material support: Bougatsos, Hartling.

Supervision: Chou, Hartling.

Conflict of Interest Disclosures: Dr Donovan reported being coauthor of the Diabetes Canada 2018 Evidence-based Pregnancy Guidelines and receiving nonfinancial support from Medtronic. No other disclosures were reported.

Funding/Support: This research was funded under contract HHS/29020150009-I, Prism Order HHS/29020150014T, from the Agency for Healthcare Research and Quality (AHRQ). US Department of Health and Human Services, under a contract to support the US Preventive Services Task Force (USPSTF).

Role of the Funder/Sponsor: Investigators worked with USPSTF members and AHRQ staff to develop the scope, analytic framework, and key questions for this review. AHRQ had no role in study selection, quality assessment, or synthesis. AHRQ staff provided project oversight, reviewed the report to ensure that the analyses met methodological standards, and distributed the draft for peer review. Otherwise, AHRQ had no role in the conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript findings. The opinions expressed in this document are those of the authors and do not reflect the official position of AHRQ or the US Department of Health and Human Services.

Additional Contributions: We thank AHRQ Medical Officers Justin Mills, MD, MPH, and Iris Mabrury-Hernandez, MD, MPH, and the USPSTF for their contributions to this project. We thank Tina Korownyk, MD (University of Alberta), for contributions during topic development, and Robin Featherstone, MLIS, and Diana Kato Lambert, MLIS (University of Alberta), for development of searches for the topic development and review stages, respectively. We also acknowledge past and current USPSTF members who contributed to topic deliberations. The USPSTF members, external reviewers, and federal partner reviewers did not receive financial compensation for their contributions.

Additional Information: A draft version of this evidence report underwent external peer review from 9 federal partners representing the Centers for Disease Control and Prevention, Indian Health Service, Office of Clinical and Preventive Services, Division of Diabetes Treatment and Prevention; National Institutes of Health, Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development; National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; Division of Health Promotion and Communication, Office of Disease Prevention and Health Promotion; Office of Research on Women’s Health; and 5 content experts (Andrew Garrison, MD, University of Utah; Joseph R. Biggio Jr, MD, Ochsner Health, New Orleans, Louisiana; Diane Farrar, PhD, Bradford Institute for Health Research, Bradford, United Kingdom; Florence M. Brown, MD, Joslin Diabetes Center, Boston, Massachusetts; Linda A. Barbour, MD, University of Colorado School of Medicine at Anschutz Medical Campus). Reviewer comments were presented to the USPSTF during its deliberation of the evidence and were considered in preparing the final evidence review.

Editorial Disclaimer: This evidence report is presented as a document in support of the accompanying USPSTF Recommendation Statement. It did not undergo additional peer review after submission to JAMA.

REFERENCES

1. American College of Obstetricians and Gynecologists. ACOG Practice Bulletin No. 190 summary: gestational diabetes mellitus. Obstet Gynecol. 2018;131(2):406-408. doi:10.1097/AOG.0000000000002498

2. American Diabetes Association. Classification and diagnosis of diabetes: standards of medical care in diabetes—2020. Diabetes Care. 2020;43 (suppl 1):S3-S31. doi:10.23736/S0149-1992-20-431-S001-00001-X

3. Casagrande SS, Linder B, Cowie CC. Prevalence of gestational diabetes and subsequent type 2 diabetes among U.S. women. Diabetes Res Clin Pract. 2018;141:200-208. doi:10.1016/j.diabres.2018.06.010

4. DeSisto CL, Kim SY, Sharma AJ. Prevalence estimates of gestational diabetes mellitus in the United States, Pregnancy Risk Assessment Monitoring System (PRAMS), 2007-2010. Prev Chronic Dis. 2014;11:E104. doi:10.5888/pchdee.2014.11.010345

5. Lavery JA, Friedman AM, Keyes KM, Wright JD, Ananth CV. Gestational diabetes in the United States: temporal changes in prevalence rates between 1979 and 2010. JOG. 2012;17(4):504-813. doi:10.1111/j.1471-0528.14236

6. Zhou T, Sun D, Li X, et al. Prevalence and trends in gestational diabetes mellitus among women in the US, 2006-2016. Diabetes. 2018;67(suppl 1):121-1-DOR. doi:10.23736/S0149-1992-20-671-S001-00019-0

7. Brown FM, Wycoff J. Application of one-step IAPDSG versus two-step diagnostic criteria for gestational diabetes in the real world: impact on health services, clinical care, and outcomes. Curr Diab Rep. 2017;17(10):85. doi:10.1007/s11892-017-0922-2

8. Metzger BE, Gabbe SG, Persson B, et al; International Association of Diabetes and Pregnancy Study Groups Consensus Panel. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care. 2010;33(3):676-682. doi:10.2337/dc09-1848

9. Pillay J, Donovan L, Guirard S, et al. Screening for Gestational Diabetes Mellitus: a Systematic Review to Update the 2014 U.S. Preventive Services Task Force Recommendation. Evidence Synthesis No. 204. Agency for Healthcare Research and Quality; 2021. AHRQ publication 21-05273-EF-1.

10. Moyer VA; US Preventive Services Task Force. Screening for gestational diabetes mellitus: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2014;160(6):414-420. doi:10.7326/M13-2905

11. Hartling L, Dryden DM, Guthrie A, et al. Screening and Diagnosing Gestational Diabetes Mellitus. Evip Rep Technol Assess No. 210. Agency for Healthcare Research and Quality; 2012. AHRQ report 12(13)-E021-EF.

12. Procedure Manual. US Preventive Services Task Force. Published May 2021. Accessed June 14, 2021. https://www.uspreventiveservicestaskforce.org/ uspstf/about-uspstf/methods-and-processes/ procedure-manual

13. Higgins J, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.10. The Cochrane Collaboration; 2011.

14. Wells G, Shea B, O’Connell N, et al. The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomized Studies in Meta-analyses. University of Ottawa Department of Epidemiology and Community Medicine; 2009.

15. Moons KGM, Woff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. Ann Intern Med. 2019;170(1):W1-W33. doi:10.7326/M18-1377

16. Whiting PF, Rutjes AW, Westwood ME, et al; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-536. doi:10.7326/0003-4819-155-8-201110180-00009

17. Berkman ND, Lohr KN, Ansari MT, et al. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. J Clin Epidemiol. 2015;68(11):1312-1324. doi:10.1016/j.jclinepi.2014.11.023

18. Basri NI, Mahdy ZA, Ahmad S, et al. The World Health Organization (WHO) versus the International Association of Diabetes and Pregnancy Study Group (IAPDSG) diagnostic criteria of gestational diabetes mellitus (GDM) and their associated maternal and neonatal outcomes. Horm Mol Biol Clin Investig. 2018;34(1). doi:10.1515/ hmbc-2017-0077

19. Harper LM, Jauk V, Longo S, Biggio JR, Szchowski JM, Tita AT. Early gestational diabetes screening in obese women: a randomized controlled trial. Am J Obstet Gynecol. 2020;222(5):495.e1-495.e8. doi:10.1016/j.ajog.2019.12.021

20. Hillier TA, Pedula KL, Ogasawara KK, et al. A pragmatic, randomized clinical trial of gestational diabetes screening. N Engl J Med. 2021;384(10):895-904. doi:10.1056/NEJMoa2026028

21. Khalifeh A, Eckler R, Felder L, Saccone G, Caissetti C, Bergella V. One-step versus two-step diagnostic testing for gestational diabetes: a randomized controlled trial. J Matern Fetal Neonatal Med. 2020;33(4):612-617. doi:10.1080/ 14767058.2019.1498480
Clinical Review & Education
US Preventive Services Task Force

22. Scifres CM, Abebe KZ, Jones KA, et al. Gestational diabetes diagnostic methods (GDM2M) pilot randomized trial. Matern Child Health J. 2015; 19(7):1472-1480. doi:10.1007/s10995-015-1651-4
23. Sevket O, Ates S, Uysal G, Molla T, Dansuk R, Kelekci S. To evaluate the prevalence and clinical outcomes using a one-step method versus a two-step method to screen gestational diabetes mellitus. J Matern Fetal Neonatal Med. 2014;27(1): 36-41. doi:10.3109/14767058.2013.799656
24. Davis EM, Abebe KZ, Simhan HM, et al. Perinatal outcomes of two screening strategies for gestational diabetes mellitus: a randomized controlled trial. Obstet Gynecol. 2021;138(1):3-10.
25. 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):269-275. doi:10.1055/s-2007-993871
26. Bonomo M, Corica D, Mion E, et al. Evaluating the therapeutic approach in pregnancies complicated by borderline glucose intolerance: a randomized clinical trial. Diabet Med. 2005;22(11): 1536-1541. doi:10.1111/j.1464-5494.2005.09690.x
27. Crowther CA, Hiller JE, McPhee AJ, Jeffries WS, Robinson JS; Australian Carbohydrate Intolerance Study in Pregnant Women (ACHOIS) Trial Group. Effect of treatment of gestational diabetes mellitus: is a diagnosis associated with an increase in maternal anxiety and stress in the short and intermediate term? Diabetes Care. 1997;20(12): 3480-3485. doi:10.2337/diacare.20.12.3480
28. Fadl HE, Gárdefos S, Hjertberg R, et al. Randomized controlled study in pregnancy on treatment of marked hyperglycemia that is short of overt diabetes. Acta Obstet Gynecol Scand. 2015;94 (11):1181-1187. doi:10.1111/aogs.12717
29. Garner P, Okun N, Keely E, et al. A randomized controlled trial of strict glycemic control and tertiary level obstetric care versus routine obstetric care in the management of gestational diabetes: a pilot study. Am J Obstet Gynecol. 1997;177(1):190-195. doi:10.1016/S0002-9378(97)00467-1
30. Hughes RCE, Rowan J, Willjman J. Prediabetes in pregnancy, can early intervention improve outcomes? a feasibility study for a randomised controlled clinical trial. BMJ Open. 2018;8(3): e018493. doi:10.1136/bmjopen-2017-018493
31. Kokanli MK, Tokmak A, Kaymak O, Cavkaytar S, Bilge Ü. The effect of treatment on pregnancy outcomes in women with one elevated oral glucose tolerance test value. Ginekol Pol. 2014;85(6):748-753.
32. Landon MB, Spong CY, Thom E, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. A multicenter, randomized trial of treatment for mild gestational diabetes. N Engl J Med. 2009;360(19):1939-1948. doi:10.1056/NEJMoa0902430
33. Osmundson SS, Norton ME, El-Sayed YY, Carter S, Faig JC, Kitzmiller JL. Early screening and treatment of women with prediabetes: a randomized controlled trial. Am J Perinatol. 2016;33(2):172-179. doi:10.1055/s-0035-1563575
34. Simmons D, Nema J, Parton C, et al. The Treatment Of Baking Gestational Diabetes Mellitus (TOBGDM) pilot randomized controlled trial. BMC Pregnancy Childbirth. 2018;18(1):151. doi:10.1186/s12884-018-1809-y
35. Yang X, Tian H, Zhang F, et al. A randomised transitional trial of lifestyle intervention using a 3-tier shared care approach on pregnancy outcomes in Chinese women with gestational diabetes mellitus but without diabetes. J Transl Med. 2014;12:290. Published correction appears in J Transl Med. 2015;13:70. doi:10.1186/1476-0729-2014-1
36. Devere R, Devere M, Akbaba E, et al. The effect of diet on pregnancy outcomes among women with abnormal glucose challenge test. Eur Rev Med Pharmacol Sci. 2013;17(9):1259-1261.
37. Vinter CA, Tanvig MH, Christensen MH, et al. Lifestyle intervention in Danish obese pregnant women with early gestational diabetes mellitus according to WHO 2013 criteria does not change pregnancy outcomes: results from the LIP (Lifestyle in Pregnancy) study. Diabetes Care. 2018;41(10): 2079-2085. doi:10.2337/dc18-0808
38. Chanprapaph P, Sutjitprat T. 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):1141-1146.
39. Hivert MF, Allard C, Menard J, Ouellet A, Ardioule J. Impact of the creation of a specialized clinic for prenatal blood sampling and follow-up care in pregnant women. J Obstet Gynecol Can. 2012;34(3):236-242. doi:10.1503/cmaj.116016(3)3594-5
40. Solomon CG, Willett WC, Rich-Edwards J, et al. Variability in diagnostic evaluation and criteria for gestational diabetes. Diabetes Care. 1996;19(12):11-16. doi:10.2337/diacare.19.12.1
41. Stacey T, Tennant P, McCowan L, et al. Gestational diabetes and the risk of late stillbirth: a case-control study from England, UK. BJOG. 2019;126(8):973-982. doi:10.1111/1460-1523.15276
42. Daniells S, Greynery BF, Davis WS, Coleman KJ, Burgess JA, Moses RG. 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): 385-389. doi:10.2337/diacare.26.2.385
43. Doughty KN, Ronnenberg AG, Reeves KW, Khan J, Sibeko L. Barriers to exclusive breastfeeding among women with gestational diabetes mellitus in the United States. J Obstet Gynecol Neonatal Nurs. 2018;47(3):301-315. doi:10.1111/jog.12005
44. Kerbel D, Glazier R, Holzapfel S, Yeung M, Lofsky S. Adverse effects of screening for gestational diabetes mellitus: a prospective cohort study in Toronto, Canada. J Med Screen. 1997;4(3):128-132. doi:10.1077/1097-0000.000616992.1207156
45. Loewenberg Weiband Y, Rausch J, Kachoria R, Gunderson EP, Oza-Frank R. Hospital supplementation differentially impacts the association between breastfeeding intention and duration among women with and without gestational diabetes mellitus history. Breastfed Med. 2017;12(6):338-344. doi:10.1089/bfm.2017.0019
46. Naylor CD, Sermer M, Chen E, et al; Toronto Breastfeeding Study Group. In-hospital breastfeeding experiences among women with gestational diabetes. Breastfeed Med. 2017;12:261-268. doi:10.1089/bfm.2016.01975
47. Braga FO, Negrapa CA, Matta MFBD, Carniero JRI, Gomes MB. Relationship between inflammatory markers, glycated hemoglobin and placental weight on fetal outcomes in women with gestational diabetes. Arch Endocrinol Metab. 2019;63(1):22-29. doi:10.20945/2339-3979.201905049
48. Patanb N, Chatri A. Time-dependent gestational diabetes screening rates. J Med Screen. 2018;25(3):121-125. doi:10.1077/1097-0000.00066
49. Braga FO, Negrapa CA, Matta MFBD, Carniero JRI, Gomes MB. Time-dependent gestational diabetes screening values. Int J Gynaecol Obstet. 1997;56(3):257-261. doi:10.1016/S0020-7292(96)00317-1
50. Chevaller N, Fénichel P, Gaume V, et al. Universal two-step screening strategy for gestational diabetes: does it work? Diabetes Metab. 2011;37(2):419-425. doi:10.1016/j.diabet.2011.01.004
Glycosylated haemoglobin for screening and diagnosis of gestational diabetes mellitus. 

Holleman S, Allen SR. Jelly beans as an alternative to a fifty-gram glucose beverage for gestational diabetes screening. 

67. 2016;6(4):e011059. doi:10.1136/bmjopen-2016-011059. 

68. Hughes RC, Moore MP, Gullam JE, Mohamed K, Rowan J. An early pregnancy HbA1c ≥ 5.3% (41 mmol/mol) is optimal for detecting diabetes and identifies women at increased risk of adverse pregnancy outcomes. 

69. Diabetes Care. 2014;37(10):2953-2959. doi:10.23736/S0149-385X.14-03839-7. 

70. Kauffman RP, Castracane VD, Peghee D, Baker TE, Van Hook JW. Detection of gestational diabetes mellitus using hemoglobin A1c: revising current screening guidelines. 

71. Cian J. Intem Med. 2019;10(1):16-24. doi:10.22038/jiim.2016.10.1.16. 

72. Poo ZX, Wright A, Ruocheng D, Singh R. Optimal first trimester HbA1c threshold to identify Singaporean women at risk of gestational diabetes mellitus and adverse pregnancy outcomes: a pilot study. Obset Med. 2019;12(2):79-84. doi:10.1177/1753495X19875984. 

73. Poosol Mol, Kangawarwari. A comparison of fasting plasma glucose and glucose challenge test for screening of gestational diabetes mellitus. 

74. J Obstet Gynaecol. 2013;33(5):447-450. doi:10.3109/01443615.2013.771556. 

75. Rajput R, Yogesh Yadav, Rajput M, Nanda S. Utility of HbA1c, for diagnosis of gestational diabetes mellitus. 

76. Diabetes Res Clin Pract. 2012;98(1):104-107. doi:10.1016/j.diabres.2012.02.018. 

77. Saccone G, Caissutti C, Khalifeh A, et al. One step versus two step approach for gestational diabetes screening: systematic review and meta-analysis of the randomized trials. 

78. J Matern Fetal Neonat Med. 2017;1-9. doi:10.1080/14767058.2017.1408068. 

79. Casey BM, Mele L, Landon MB, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. Does maternal body mass index influence treatment effect in women with mild gestational diabetes? 

80. Am J Perinatol. 2015;32(1):93-100. doi:10.1055/s-0034-1374815. 

81. Casey BM, Rice MM, Landon MB, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. Effect of treatment of mild gestational diabetes on long-term maternal outcomes. 

82. Harper LM, Mele L, Landon MB, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) Maternal-Fetal Medicine Units (MFMU) Network. Carpenter-Coustan compared with National Diabetes Data Group criteria for diagnosing gestational diabetes. 

83. Obstet Gynecol. 2016;127(5):893-898. doi:10.1097/AOG.0000000000001383.
98. Landon MB, Rice MM, Varner MW, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal-Fetal Medicine Units (MFMU) Network. Mild gestational diabetes mellitus and long-term child health. *Diabetes Care*. 2015;38(3):445-452. doi:10.2337/dc14-2159

99. Palatnik A, Mele L, Landon MB, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. Timing of treatment initiation for mild gestational diabetes mellitus and perinatal outcomes. *Am J Obstet Gynecol*. 2015;213(4):560.e1-560.e8. doi:10.1016/j.ajog.2015.06.022

100. Berggren EK, Mele L, Landon MB, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) Maternal-Fetal Medicine Units (MFMU) Network. Perinatal outcomes in Hispanic and non-Hispanic white women with mild gestational diabetes. *Obstet Gynecol*. 2012;120(5):1099-1104. doi:10.1097/AOG.0b013e31827049a5

101. Gillman MW, Oakey H, Baghurst PA, Vollmer RE, Robinson JS, Crowther CA. Effect of treatment of gestational diabetes mellitus on obesity in the next generation. *Diabetes Care*. 2010;33(5):964-968. doi:10.2337/dc09-1810

102. Malcolm JC, Lawson ML, Gaboury I, Lough G, Keely E. Glucose tolerance of offspring of mother with gestational diabetes mellitus in a low-risk population. *Diabet Med*. 2006;23(5):565-570. doi:10.1111/j.1464-5491.2006.01840.x

103. The IDEAL Randomised Controlled Trial [ACTRN12607000174482]. Australian New Zealand Clinical Trials Registry. Updated October 1, 2017. Accessed March 4, 2021. https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=81643

104. Efficacy of Treatment for Gestational Diabetes Diagnosed by the IADPSG Criteria [NCT02728758]. ClinicalTrials.gov. Updated August 13, 2019. Accessed March 4, 2021. https://clinicaltrials.gov/ct2/show/NCT02708758

105. Moher D, Pham B, Klassen TP, et al. What contributions do languages other than English make on the results of meta-analyses? *J Clin Epidemiol*. 2000;53(9):964-972. doi:10.1016/S0895-4356(00)00188-8

106. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002. doi:10.1136/bmj.d4002

107. Cornell JE, Mulrow CD, Localio R, et al. Random-effects meta-analysis of inconsistent effects: a time for change. *Ann Intern Med*. 2014;160(4):267-270. doi:10.7326/M13-2886

108. Lee D, Booth GL, Ray JG, Ling V, Feig DS. Undiagnosed type 2 diabetes during pregnancy is associated with increased perinatal mortality: a large population-based cohort study in Ontario, Canada. *Diabet Med*. 2020;37(10):1696-1704. doi:10.1111/dme.14250

109. Simmons D, Hague WM, Teede HJ, et al. Hyperglycaemia in early pregnancy: the Treatment Of Booking Gestational Diabetes Mellitus (TOBOGM) study: a randomised controlled trial. *Med J Aust*. 2018;209(9):405-406. doi:10.5694/mja17.01129

110. Farrar D, Simmonds M, Bryant M, et al. Hyperglycaemia and risk of adverse perinatal outcomes: systematic review and meta-analysis. *BMJ*. 2016;354:i4694. doi:10.1136/bmj.i4694