Screening for Gestational Diabetes: A Summary of the Evidence for the U.S. Preventive Services Task Force

By

Seth C. Brody, M.D.

A Master's paper submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Public Health in the School of Public Health, Public Health Leadership Program

Approved by:

William Miller, Advisor

Russ Harris, Content Reader

Date
Screening for Gestational Diabetes: A Summary of the Evidence for the U.S. Preventive Services Task Force

By

Seth C. Brody, M.D.

A Master's paper submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Public Health in the School of Public Health, Public Health Leadership Program

Approved by:

William Miller, Advisor

Russ Harris, Content Reader

Date
Screening for Gestational Diabetes: A Summary of the Evidence for the U.S. Preventive Services Task Force

Seth C. Brody, M.D., Russell Harris, M.D., M.P.H., Kathleen Lohr, Ph.D.

Epidemiology

Gestational diabetes mellitus (GDM) is defined as glucose intolerance with the onset or first detection during pregnancy. About 135,000 cases of GDM are diagnosed annually in the United States. Important risk factors include higher maternal age, family history of diabetes, and increased pregravid body mass index (BMI). The prevalence of GDM in low-risk populations ranges from 1.4% to 2.8%; in high-risk populations, prevalence ranges from 3.3% to 6.1%.

Markedly elevated maternal glucose levels most often occur in women with pregestational diabetes. Pregnant women with pregestational diabetes are at higher risk for multiple complications affecting both the mother and the fetus than those women without diabetes. Current therapy improves outcomes for both mother and neonate.

The additional risk for adverse health outcomes attributable to the milder degrees of maternal hyperglycemia associated with GDM and the magnitude of the benefit from treating that risk are less certain. No well-designed and conducted randomized controlled trial (RCT) of screening for GDM has been completed, and thus the evidence for screening is indirect.

National groups disagree about whether to recommend screening for GDM. Despite no strong recommendations in favor of universal screening from the American College of Obstetricians and Gynecologists (ACOG), 94% of Fellows in office-based practices reported performing universal screening for GDM in 1996. Fellows performed this screening even though ACOG acknowledged the weakness in the evidence in both 1994 and 2000.

With continued controversy around the advisability of GDM screening, the RTI–University of North Carolina Evidence-based Practice Center (RTI–UNC EPC) conducted a systematic evidence review to assist the U.S. Preventive Services Task Force (USPSTF) in reconsidering its 1996 review, which found insufficient evidence to recommend screening. We restricted this review to screening for GDM after 24 weeks' gestation, thus excluding both women with known pregestational diabetes and those who are discovered by symptoms earlier in pregnancy.
Screening for Gestational Diabetes: A Summary of the Evidence for the U.S. Preventive Services Task Force

Seth C. Brody, M.D., Russell Harris, M.D., M.P.H., Kathleen Lohr, Ph.D.

Epidemiology

Gestational diabetes mellitus (GDM) is defined as glucose intolerance with the onset or first detection during pregnancy. About 135,000 cases of GDM are diagnosed annually in the United States. Important risk factors include higher maternal age, family history of diabetes, and increased pregravid body mass index (BMI). The prevalence of GDM in low-risk populations ranges from 1.4% to 2.8%; in high-risk populations, prevalence ranges from 3.3% to 6.1%.

Markedly elevated maternal glucose levels most often occur in women with pregestational diabetes. Pregnant women with pregestational diabetes are at higher risk for multiple complications affecting both the mother and the fetus than those women without diabetes. Current therapy improves outcomes for both mother and neonate.

The additional risk for adverse health outcomes attributable to the milder degrees of maternal hyperglycemia associated with GDM and the magnitude of the benefit from treating that risk are less certain. No well-designed and conducted randomized controlled trial (RCT) of screening for GDM has been completed, and thus the evidence for screening is indirect.

National groups disagree about whether to recommend screening for GDM. Despite no strong recommendations in favor of universal screening from the American College of Obstetricians and Gynecologists (ACOG), 94% of Fellows in office-based practices reported performing universal screening for GDM in 1996.

With continued controversy around the advisability of GDM screening, the RTI–University of North Carolina Evidence-based Practice Center (RTI–UNC EPC) conducted a systematic evidence review to assist the U.S. Preventive Services Task Force (USPSTF) in reconsidering its 1996 review, which found insufficient evidence to recommend screening. We restricted this review to screening for GDM after 24 weeks' gestation, thus excluding both women with known pregestational diabetes and those who are discovered by symptoms earlier in pregnancy.
Materials and Methods

Sources
Our review of the literature was guided by key questions and inclusion criteria we developed relevant to the issue of screening for GDM (Table 1). We required RCTs for direct evidence of the efficacy of treatment and the harms associated with treatment. We examined the critical literature from the 1996 USPSTF review and searched MEDLINE and the Cochrane Library for reviews and relevant studies published in English between January 1, 1994 and August 30, 2002. We supplemented this search by examining the reference lists of pertinent articles and by contacting experts. We also conducted focused searches of MEDLINE from 1966 through 1994 to identify older articles of interest.

Study Selection
All searches began by exploding the term "diabetes, gestational" and then proceeded by adding further terms. We retrieved the full text of all articles we thought were potentially eligible. Two reviewers examined each article for eligibility. A single reviewer abstracted relevant data from the included articles; a second reviewer checked the abstractions.

We abstracted all included articles, entered the data into evidence tables, graded the quality of all articles according to USPSTF criteria, and resolved disagreements by discussion. We synthesized the evidence into a systematic evidence review; this was subjected to extensive external peer review and revised as appropriate. The final systematic evidence review, including the evidence

| Table 1. Inclusion criteria, search strategy, and results of searches for 6 key questions |
|---------------------------------|---------------------------------|------------------|
| **Key question** | **Inclusion criteria** | **Number of articles meeting criteria** |
| 1. Screening efficacy for maternal and fetal health outcomes | RCT, Screening for GDM, Maternal or infant health outcomes | 0 |
| 2. Adverse health outcomes of untreated GDM | Any research design, Screening for GDM, Maternal or infant health outcomes | 9 |
| 3. Accuracy and reliability of screening tests | Screening test for GDM, Data available to calculate sensitivity and specificity, Criterion standard used | 13 |
| 4. Treatment for GDM: | | 9 |
| - Glycemic control | RCT, Glycemic control, Health outcomes | |
| - Antepartum surveillance | RCT, Surveillance or antepartum, Health outcomes | 5 |
| 5. Harms of screening and treatment | Any research design, Any harm associated with screening or treatment of GDM | 9 |
| 6. Costs, efficiency of screening | Any research design, Costs, efficiency of screening for GDM | 7 |

Note: All searches started with exploding "diabetes, gestational."
GDM indicates gestational diabetes mellitus; RCT, randomized clinical trial.
tables, is available on the Agency for Healthcare Research and Quality (AHRQ) Web site (www.preventiveservices.ahrq.gov). This article summarizes the evidence from that review.

Results

For the USPSTF to recommend screening for GDM, it must have either direct evidence from a randomized controlled trial (RCT) of screening or indirect evidence that establishes a complete linkage between screening and improved health outcomes. We found no well-conducted RCT that provides direct evidence for the health benefits of screening for GDM. Given this, the USPSTF requires adequate evidence that: (1) untreated GDM causes substantial maternal and/or neonatal adverse health outcomes; (2) available screening tests accurately and efficiently detect GDM; and (3) available treatments improve health outcomes, with a magnitude that clearly justifies the harms and effort of screening and treatment. These issues will be examined in the sections that follow.

What Adverse Health Outcomes Occur with Untreated GDM?

Determining the existence and magnitude of a causal association between various degrees of GDM and adverse health outcomes is complex. We have only older studies of untreated GDM, at a time when obstetric practice differed from current practice, or more recent studies in which women received some treatment for GDM. Another problem with many studies is that they consider GDM as a dichotomous variable, yet we know that the risk for adverse health outcomes increases with the degree of hyperglycemia among women with GDM; the impact of hyperglycemia on adverse maternal and neonatal health outcomes is probably continuous. Few studies, however, stratify the risks of GDM by severity of hyperglycemia.

Offspring Health Outcomes

Because the literature is scant and mixed about whether untreated GDM, given optimal obstetric care today, is associated with increased perinatal mortality, the extent to which GDM is truly associated with perinatal mortality remains unclear.

Macrosomia is an intermediate outcome of GDM. Three recent studies of untreated women with GDM found that the percentage of infants with macrosomia weighing more than 4,000 grams was between about 17% and 29%; the percentage in the general population is about 10%. Most infants with macrosomia are born to women without GDM; maternal obesity is an important potential confounding factor associated with both GDM and (independently) with macrosomia.

Important adverse neonatal health outcomes linked to macrosomia are brachial plexus injury and clavicular fracture. The best (although minimal) data on untreated women with GDM compared with the non-GDM population reveal no difference in the rate of infant brachial plexus injury or clavicular fracture. Recent data suggest that women treated for GDM with more severe degrees of hyperglycemia may have a 2% absolute increase in having their infants develop a brachial plexus injury and a 6% increase in having their infants develop a clavicular fracture. While these adverse health outcomes are of concern, the best studies show that 80% to 90% of brachial plexus injuries resolve by one year of life, and more than 95% of clavicular fractures heal within a few months without residual problems.

GDM may also be a risk factor for neonatal hypoglycemia. Studies among untreated and treated women with GDM have found higher rates of neonatal hypoglycemia among untreated women with GDM. The magnitude of clinically important neonatal hypoglycemia is less clear. Also not clear is whether increased surveillance of infants whose mothers have GDM contributes to the increased finding of hypoglycemia in their infants.

Likewise, the evidence is limited and unclear as to whether GDM is associated with preterm birth or neonatal hyperbilirubinemia, hypocalcemia, or polycythemia. Because of limited evidence and the increased surveillance given to infants of women with GDM, the magnitude of any
associated adverse health effects is uncertain but is likely to be small.

Some have suggested that the diagnosis of maternal GDM may have long-term implications for the offspring, such as an increased risk for impaired glucose tolerance, childhood obesity, and neuropsychological disturbances. No large observational study has followed a group of children whose mothers have GDM and a comparison group whose mothers do not have GDM long enough to demonstrate whether any of these hypotheses are correct.4

Maternal Health Outcomes

The diagnosis of GDM can also increase adverse health outcomes for the mother during her pregnancy. Limited data since 1980 reveal total cesarean delivery rates of 22% to 30% for unrecognized or untreated women with GDM, compared with a rate of about 17% for women without GDM. Although the overall literature suggests an association between GDM and higher cesarean delivery rates, some studies are limited by a lack of adjustment for maternal obesity and by the impact of the diagnosis of GDM on clinical decision-making.

Limited evidence is available on the rate of third- or fourth-degree lacerations in women with GDM. Some studies have suggested an increase, but the only study that found a substantial percentage of women with untreated GDM who had such lacerations included only 16 subjects. Another study found equally low rates among women with GDM and women without GDM.

Overall, observational studies have shown mixed results and are inconclusive as to whether women with GDM have a higher risk for pre-eclampsia than women without GDM. Recent data from untreated women with GDM reveal a rate of pre-eclampsia (about 9%) that is similar to that for treated women and women without GDM.

Mothers identified as having gestational diabetes also have a higher risk for developing type 2 diabetes in the years after delivery. Studies investigating the rate of development of type 2 diabetes after the onset of gestational diabetes have suffered from low participation rates, retrospective design, short follow-up, and variation in definition of both GDM and new diabetes. Although nearly all studies show that women who have GDM face some increased risk for developing diabetes, the degree of risk elevation they experience and the degree of glucose abnormality they develop are uncertain. Further, the added benefit of early detection of diabetes in young women with few cardiovascular risk factors is uncertain.

How Accurate and Reliable Are Screening Tests for GDM?

Reference Diagnostic Test

Before we can determine the accuracy of a screening test, we need a reference diagnostic test for comparison. Unfortunately, no universally agreed on reference test for the diagnosis of GDM exists.

Three competing criteria for diagnostic glucose tolerance tests (GTT) are available (Table 2). Criteria from the World Health Organization (WHO) label twice as many women with GDM as do criteria from the National Diabetes Data Group (NDDG). Criteria from the American Diabetes Association (ADA) give an intermediate prevalence.

Abnormal values on any of the 3 tests are predictive of fetal macrosomia. This association is diminished or eliminated when adjustments are made for such potential confounders as prepregnancy weight, age, parity, and race.

The reliability of any oral GTT is open to question. In 1 of the few studies on this issue, Harlaas et al found that 23% of 64 unscreened pregnant women who had had a positive screening test for GDM had inconsistent results between two 100-gram oral GTTs performed 1 week apart. Other studies have also raised concerns about the reproducibility of the oral GTT in nonpregnant groups.
Two or more criteria must be met or exceeded for a positive diagnosis. One or more criteria must be met or exceeded for a positive diagnosis.

Note: Double dash (—) indicates glucose levels not used for the test indicated.

Screening Tests. The thresholds for the reference diagnostic tests do not clearly distinguish women at high risk from women at low risk for adverse maternal or neonatal health outcomes from GDM. Thus, we can evaluate screening tests only against imperfect standards. Most studies on GDM screening strategies compare the results of 1 test with the results of another test rather than examining how the test predicts adverse health outcomes. Some studies assess the association of the test with intermediate outcomes such as macrosomia rather than health outcomes such as brachial plexus injury.

In the United States, the 50-gram, 1-hour glucose challenge test (GCT) is most commonly used for screening (Table 2). Two groups have proposed different threshold criteria to define a positive screening test. If the GCT glucose value is above either 130 mg/dL or 140 mg/dL, then the patient is usually given the 100-gram GTT for diagnosis. Using the 130 mg/dL threshold, the GCT is positive for 20% to 25% of all pregnant women, including 90% of women with GDM. Using the 140 mg/dL threshold, the GCT is positive for 14% to 18% of all pregnant women, including about 80% of women with GDM.

In the general population, false-positive results for the GCT are common. Fewer than 1 in 5 women with a positive GCT will meet criteria for GDM on a full 100-gram GTT. The reliability of the GCT is also problematic.

In many countries outside North America, clinicians use the WHO screening approach: the 75-gram 2-hour oral GTT as a single-step screening and diagnostic test. As noted above, this approach identifies at least twice as many women as having GDM as the two-step approach, although the evidence is sparse about whether the one-step test is more or less predictive of adverse health outcomes than the two-step approach.

Because glucose intolerance increases during pregnancy, screening for GDM is most commonly conducted during the 24th to 28th week of gestation. However, this timing is not based on any evidence that this is the optimal time to identify women who would benefit most from treatment. Determining the best time to screen involves examining the trade-off between the potential benefits of early screening (ie, finding fewer women at higher risk and treating them for a longer time) and the potential benefits of later screening (ie, finding a larger number of women at lower risk and treating them for a shorter time). We found no study on this issue.

One suggested approach to improve the efficiency of screening for GDM is to restrict screening to women at higher risk ("selective screening") rather than screening all women.
In the most detailed study of selective screening strategies, Naylor et al developed a scoring system that excluded nearly 35% of women from screening and actually detected more cases of GDM than universal screening. In summary, the evidence is unclear about the optimal screening and reference diagnostic test for screening for GDM.

**Does Treatment for GDM Improve Health Outcomes?**

**Glycemic Control**

Three factors are important in considering studies that evaluate the impact of tight glycemic control on health outcomes for women with GDM. The first is the degree of hyperglycemia in study participants. As the risk for at least some adverse health events increases with an increasing degree of hyperglycemia, the potential absolute risk reduction may be larger with higher glycemic levels. More than 70% of women diagnosed with GDM have mild hyperglycemia and are usually treated with diet alone.

The second important factor is the degree of separation of glycemic control between treatment groups. If intensive treatment does not produce a reasonable reduction in glycemic level compared with conventional treatment (or no treatment), the hypothesis of improved glycemic control leading to better health outcomes cannot be tested.

The third factor in considering these studies is assessment of outcomes: which ones to assess and how to assess them. Most of these studies focused on intermediate outcomes such as fetal macrosomia or chemical findings such as neonatal hypoglycemia. Intermediate outcomes are useful only insofar as they predict important health outcomes that people care about. In the case of fetal macrosomia, an intermediate outcome, only a small percentage of these cases lead to maternal or neonatal trauma. In the case of chemical findings (eg, glucose or bilirubin level), few studies reported the percentage of abnormalities that required treatment; no study was clearly reassuring that differences attributed to improved glycemic control were not associated instead with more intense surveillance of infants born to GDM mothers. Finally, because few of these studies masked the obstetricians, interventions or outcomes that depend on clinician judgment (eg, cesarean delivery rates) could be biased by knowledge of GDM status.

Table 3 records data from 9 RCTs examining the impact of therapy on a variety of outcomes. The first 4 RCTs are of women with mild hyperglycemia and the last 5 are of women with severe or very severe hyperglycemia.

**Mild hyperglycemia.** Few studies have examined the effectiveness of intensive compared with less intensive glycemic control among women with GDM who have mild hyperglycemia. An overview of 4 trials that included 612 women with mild hyperglycemia found no difference in adverse health outcomes between the women treated with a modified diet and the women receiving no therapy.

The Li et al RCT made a similar comparison and had similar findings.

Three RCTs compared intensive with less intensive glycemic control (achieving some glycemic separation) among women with GDM who had varying degrees of hyperglycemia but a low mean entry fasting plasma glucose (FPG) or mean hemoglobin A1c (HbA1c). Two studies found statistically significant improvements in intermediate outcomes for those women undergoing intensive glycemic control (eg, fewer large for gestational age [LGA] infants; lower incidence of neonatal hypocalcemia); no study found clear differences in health outcomes between glycemic control groups.

**Severe hyperglycemia.** Four RCTs examined tight and less tight glycemic control among women with GDM at more severe hyperglycemic levels (Table 3). Of these trials, 3 achieved either small or no difference in glycemic control between groups and found no difference in major outcomes. One trial found a small absolute reduction in chemical abnormalities in the neonate; another found a reduction in cesarean deliveries that was not explained by fetal size.
One study achieved a larger glycemic separation between groups (difference in mean glucose, 24 mg/dL). The infants of less intensively treated women had a higher mean birth weight plus higher rates of neonatal hypoglycemia and polycythemia. These differences were small and of uncertain clinical importance.

Finally, de Veciana et al compared tight with less tight control among women with very severe hyperglycemia, some of whom likely had frank diabetes (Table 3). They also achieved the largest separation in glycemic control (HbA1c difference of 1.6%) and found some of the larger reductions in fetal macrosomia and neonatal hypoglycemia. Given the study population, however, this trial may have little relevance for the great majority of women detected with GDM.

A major issue in all of these trials is that they have too few participants to be able to detect small differences among treatment groups in such uncommon adverse health outcomes as perinatal mortality and brachial plexus injury. They have even less power to determine whether the health benefit is different for women with GDM who have severe hyperglycemia compared with those who have mild hyperglycemia. They provide insufficient evidence to confirm or refute the hypothesis that glycemic control improves health outcomes for women with GDM.

Several observational studies without randomized controls have suggested improved intermediate or health outcomes with more intensive treatment of women with GDM. The weakness in these studies is that women in the treatment groups differ from women in the control groups in multiple ways (some known and some unknown) other than glycemic control; most of the known factors are also associated with health outcomes. Thus, observed improvements in health outcomes may be attributable to factors other than glycemic control.

In summary, although insulin therapy decreases the incidence of fetal macrosomia for those women with more severe degrees of hyperglycemia, the magnitude of any effect on maternal and neonatal health outcomes is not clear. The evidence is insufficient to determine the magnitude of health benefit of tight glycemic control among the large number of women with GDM at milder degrees of hyperglycemia.

Antepartum Surveillance

Various approaches to antepartum surveillance might improve health outcomes among women with GDM. For non-stress testing (NST) or biophysical profile (BPP) to constitute a rationale for GDM screening, evidence would need to show that the use of these tests reduces stillbirth among women with GDM who have no other indication for these tests. This would require a large RCT, as most women with GDM have a low risk for having a stillbirth. No completed study of women with GDM has examined health outcomes among groups randomized to receive or not receive NST or BPP. Observational studies have found that using NSTs or BPPs in women with GDM is associated with lower fetal death can be attributed to the additional procedures. NSTs or BPPs have high false-positive rates, and they lead to interventions that may, on occasion, be unnecessary.

Ultrasound assessment of abdominal circumference to allow improved targeting of insulin therapy in order to decrease fetal macrosomia and birth trauma has been studied. Three RCTs have enrolled women with hyperglycemia into insulin therapy triggered by ultrasound abdominal circumference. These studies have not found any important differences in health outcomes; all 3 lacked power to detect differences in health outcomes and in none were the obstetricians masked to the intervention group.

What are the Harms and Costs of Screening and Treatment?

Precise evidence on the harms and costs of screening for GDM and early treatment is lacking. Although not well documented, the potential for adverse psychological effects from screening is real; in the general population, more than 80% of all positive GCT screening tests are false positives.
**Table 3. Randomized controlled trials of treatment for gestational diabetes mellitus**

<table>
<thead>
<tr>
<th>Author, year, total N</th>
<th>Glycemic level of participants</th>
<th>Randomization</th>
<th>GDM diagnosis inclusion</th>
<th>Glycemic separation during study</th>
<th>% Stillbirth (stat sig)</th>
<th>Birth weight &gt;4,000 g (stat sig)</th>
<th>% Large for gestational age (stat sig)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li et al., 1987 N=153</td>
<td>Low</td>
<td>A: Control: No treatment (n = 73)</td>
<td>GDM by NDDG65 criteria and normal or impaired glucose tolerance by WHO64 criteria</td>
<td>NR</td>
<td>NR</td>
<td>A: 7</td>
<td>A: 22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B: Treatment; diet, monitoring (n = 85)</td>
<td></td>
<td></td>
<td></td>
<td>B: 4 (NS)</td>
<td>B: 18 (NS)</td>
</tr>
<tr>
<td>Buchanan et al., 1994</td>
<td>N = 59 Low</td>
<td>A: Diet (n = 29)</td>
<td>GDM and fasting blood glucose &lt;105 mg/dL; fetal ultrasound AC = 75th percentile</td>
<td>5.4–10.8 mg/dL mean glucose difference in mixed-meal tolerance test.</td>
<td>NR</td>
<td>NR</td>
<td>A: 45</td>
</tr>
<tr>
<td>Garner et al., 1997</td>
<td>N = 200 Low</td>
<td>A: Routine care (n = 153)</td>
<td>Gestational diagnosis and treated with insulin if fasting blood glucose &gt;140 mg/dL or 1-hr postprandial value &gt;200 mg/dL (n = 149)</td>
<td>Lower in treated group by 5–9 mg/dL 1 hr postprandial</td>
<td>A: 0</td>
<td>A: 18.7 (NS)</td>
<td>B: 0 (NS)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B: Strict glycemic control and tertiary care (n = 149)</td>
<td></td>
<td></td>
<td></td>
<td>B: 16.1 (NS)</td>
<td>B: 16.1 (NS)</td>
</tr>
<tr>
<td>Bancroft et al., 2000</td>
<td>N = 68 Low</td>
<td>A: Diet and no diabetic monitoring (n = 35)</td>
<td>WHO64 criteria: Fasting blood glucose &lt;126 mg/dL; 2-hr 75 g = 140–200 mg/dL</td>
<td>HbA1c: 0.2%–0.7% difference</td>
<td>A: 0</td>
<td>NR</td>
<td>A: 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B: Diet and intensive diabetic monitoring (n = 32)</td>
<td></td>
<td></td>
<td></td>
<td>B: 0 (NS)</td>
<td>B: 8 (NS)</td>
</tr>
</tbody>
</table>

**Note:** AC indicates abdominal circumference; CPD, cephalopelvic disproportion; GDM, gestational diabetes mellitus; NDDG, National Diabetes Data Group; NR, not reported; NS, not statistically significant; stat sig, statistical significance; WHO, World Health Organization.
Table 3. Randomized controlled trials of treatment for gestational diabetes mellitus (continued)

<table>
<thead>
<tr>
<th>% Brachial plexus injury (stat sig)</th>
<th>% Clavicular fracture (stat sig)</th>
<th>% Hypoglycemia (stat sig)</th>
<th>% Hyperbilirubinemia (stat sig)</th>
<th>% Hypocalcemia (stat sig)</th>
<th>% Total Cesaerean delivery (stat sig)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: 0 (NS)</td>
<td>A: 0</td>
<td>No difference</td>
<td>NR</td>
<td>NR</td>
<td>A: 26</td>
</tr>
<tr>
<td>B: 0 (NS)</td>
<td>B: 0</td>
<td></td>
<td>NR</td>
<td>NR</td>
<td>B: 27</td>
</tr>
<tr>
<td>A: 0 (NS)</td>
<td>A: 0</td>
<td>A: 18</td>
<td>NR</td>
<td>NR</td>
<td>A: 14–21</td>
</tr>
<tr>
<td>B: 0 (NS)</td>
<td>B: 0</td>
<td>B: 14</td>
<td>(NS)</td>
<td>(NS)</td>
<td>B: 43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(P &lt; 0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A: 0 (NS)</td>
<td>A: 0</td>
<td>A: 8.7</td>
<td>A: 6.8</td>
<td>A: 30</td>
<td>A: 18.6</td>
</tr>
<tr>
<td>B: 0 (NS)</td>
<td>B: 0</td>
<td>B: 14.1</td>
<td>B: 5.4</td>
<td>B: 40.9</td>
<td>B: 20.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(P &lt; 0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>A: 31 (NS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B: 31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

continued on page 10
<table>
<thead>
<tr>
<th>Author, year, total N</th>
<th>Glycemic level of participants</th>
<th>Randomization</th>
<th>GDM diagnosis and inclusion</th>
<th>Glycemic separation during study</th>
<th>% Stillbirth (stat sig)</th>
<th>Birth weight &gt;4000 g (stat sig)</th>
<th>% Large for gestational age (stat sig)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parson et al, 1986; 60</td>
<td>High</td>
<td>A: Diet, add insulin for high glucose (n = 105)</td>
<td>Impaired glucose tolerance</td>
<td>No difference</td>
<td>A: 0 (NS)</td>
<td>B: 0 (NS)</td>
<td>A: 13 (NS)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B: Diet and insulin (n = 97)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Langer et al, 1989; 2</td>
<td>High</td>
<td>A: Controls: no treatment (n = 149)</td>
<td>NDDG65 criteria. One on 3-hr glucose tolerance test</td>
<td>24 mg/dL difference in blood glucose</td>
<td>NR</td>
<td>NR</td>
<td>A: 26 (P = 0.05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B: Treatment: diet and/or insulin (n = 126)</td>
<td>B: Goal of glucose = 95 mg/dL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nachum et al, 1999; 5</td>
<td>High</td>
<td>A: Diet and twice daily insulin (n = 146)</td>
<td>NDDG65 criteria.</td>
<td>3.4 mg/dL difference in mean blood glucose; 0.3%</td>
<td>A: 0.7 (NS)</td>
<td>B: 0 (NS)</td>
<td>A: 30 (NS)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B: Diet and 4 times daily insulin (n = 138)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kjos et al, 2001; 2</td>
<td>High</td>
<td>A: Standard: insulin (n = 48)</td>
<td>Fasting plasma glucose &gt;105 and &lt;120 mg/dL</td>
<td>Mean fasting plasma glucose: 84.9 (A) 3.2 mg/dL difference</td>
<td>No difference only one reported</td>
<td>A: 4.2 (NS)</td>
<td>B: 6.3 (NS)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B: Experimental insulin only if fetal AC is ≥70th percentile (n = 48)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Veciana et al, 1995; 9</td>
<td>Very high</td>
<td>A: Preprandial monitoring (n = 33)</td>
<td>NDDG65 criteria in HbA1c</td>
<td>1.6 difference in HbA1c</td>
<td>A: 3 (NS)</td>
<td>B: 0 (NS)</td>
<td>A: 42 (P = 0.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B: Postprandial monitoring (n = 33)</td>
<td>Fasting plasma glucose &gt;140 mg/dL or 1-hr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: AC indicates abdominal circumference; CPO, cephalopelvic disproportion; GDM, gestational diabetes mellitus; NDDG, National Diabetes Data Group; NR, not reported; NS, not statistically significant; stat sig, statistical significance; WHO, World Health Organization.
<table>
<thead>
<tr>
<th>Neonatal outcomes</th>
<th>% Brachial plexus injury (stat sig)</th>
<th>% Clavicular fracture (stat sig)</th>
<th>% Hypoglycemia (stat sig)</th>
<th>% Hyperbilirubinemia (stat sig)</th>
<th>% Hypocalcemia (stat sig)</th>
<th>% Total Cesaearian delivery (stat sig)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>A: 0</td>
<td>B: 20 (NS)</td>
<td>A: 6.7 (NS)</td>
<td>NR</td>
</tr>
<tr>
<td>NR</td>
<td>NR</td>
<td>A: 13 (P &lt; 0.02)</td>
<td>A: 14 (NS)</td>
<td>NR</td>
<td>A: 11 (NS)</td>
<td>B: 10 (NS)</td>
</tr>
<tr>
<td>A: 2.2 (NS)</td>
<td>A: 0 (NS)</td>
<td>A: 5.9 (P &lt; 0.02)</td>
<td>A: 21 (NS)</td>
<td>A: 0 (NS)</td>
<td>A: 28 (NS)</td>
<td>No difference (small number)</td>
</tr>
<tr>
<td>B: 1.4 (NS)</td>
<td>B: 0 (NS)</td>
<td>B: 0.7 (NS)</td>
<td>B: 11 (NS)</td>
<td>B: 7 (NS)</td>
<td>B: 28 (NS)</td>
<td>No difference (small number)</td>
</tr>
<tr>
<td>No difference (small number)</td>
<td>No difference (small number)</td>
<td>A: 10.4 (NS)</td>
<td>A: 2 (NS)</td>
<td>NR</td>
<td>A: 14.6 (P = 0.03)</td>
<td>(Greater % of women with previous cesaearean delivery in Group B)</td>
</tr>
<tr>
<td>A: 0 (NS)</td>
<td>A: 3 (NS)</td>
<td>A: 21 (P &lt; 0.05)</td>
<td>A: 12 (NS)</td>
<td>NR</td>
<td>A: 39 (CPD: 3)</td>
<td>B: 24 (CPD: 12) (P &lt; 0.04)</td>
</tr>
</tbody>
</table>
Limited and mixed data suggest that labeling may negatively influence women's perceptions of their health during pregnancy\textsuperscript{97-100} and that women diagnosed with GDM may have long-term changes in their perception of their own health.\textsuperscript{98,101} The long-term impact of these changes in perception of health is unclear.

Identification of GDM may also needlessly increase the use of NSTs or BPPs (triggering unnecessary interventions due to false positives) and rates of cesarean delivery (because of a lowered intervention threshold).\textsuperscript{24,70} Furthermore, additional tests and procedures increase the cost of screening programs. Because of the lack of evidence, the magnitude of other potential harms of aggressive glycemic-lowering therapy, such as increased maternal starvation ketosis and infants who are small for gestational age, is difficult to quantify.\textsuperscript{18,102}

As the effectiveness of screening in improving health outcomes is uncertain, so the cost-effectiveness cannot be calculated with any precision. We do not have good information about the differences in health care costs between screened and unscreened women.

### Discussion

Maternal and neonatal morbidity increase with increasing levels of maternal hyperglycemia. Screening and intensive treatment for GDM aim to reduce this morbidity. Various screening strategies can detect women with different degrees of hyperglycemia, but the threshold at which health outcomes begin to deteriorate to a clinically important degree is uncertain.

The magnitude of any benefit of intensive treatment at the various levels of hyperglycemia associated with GDM is also uncertain, but it is likely to be small among the many women with mild hyperglycemia. For women with GDM who have more severe hyperglycemia, intensive treatment is likely to reduce macrosomia. The extent to which this translates into reductions in birth trauma is uncertain but probably substantially less than reductions in macrosomia.

The evidence about the health outcomes of intensive treatment of women with GDM at various levels of maternal hyperglycemia is indirect. It is also limited by a small number of studies, small number of participants, lack of masking of obstetrical care, lack of control for important confounders, and lack of emphasis on health outcomes rather than intermediate outcomes.

By making various assumptions, we can calculate the number needed to screen (NNS) to prevent various adverse health outcomes. Take, for example, the number of women needed to screen to prevent 1 case of brachial plexus injury (Table 4). Assume that 4% of pregnant women have GDM,\textsuperscript{,} that 30% of them will have a high enough glycemic level to require insulin,\textsuperscript{79} and that, among these women, the macrosomia rate is reduced to the degree seen in the most positive study.\textsuperscript{99} The NNS to prevent one brachial plexus injury is about 8,900 (Case 1, Table 4).\textsuperscript{32,33,103} If we make more generous assumptions, the NNS becomes 3,300 (best case scenario, shown in Case 3 and footnote, Table 4). Assumptions including a lesser reduction in macrosomia, accounting for cesarean delivery rates, or using an outcome of permanent brachial plexus injury, would give much higher NNS estimates.\textsuperscript{2}

One potential benefit of detecting women with GDM is the knowledge that they have a higher risk for developing type 2 diabetes. The extent to which this information can lead to a health benefit for younger women with few cardiovascular risk factors, however, is uncertain.\textsuperscript{63}

The evidence concerning the harms and costs of screening and intensive treatment is even more limited than the evidence about benefits, but several harms are of concern. Many women may suffer anxiety of uncertain duration because of a false-positive screening test. Labeling women with GDM as having an increased risk for future GDM and type 2 diabetes may have psychological implications. Detection of GDM may increase the probability of cesarean delivery; multiple antenatal tests may increase the probability of a false-positive test leading to unnecessary procedures. Costs may be increased with little health benefit for many women, especially those many women with mild hyperglycemia.

It is difficult to see how the issue of screening for GDM can be clarified without large RCTs with
### Table 4. Number needed to screen (NNS) for gestational diabetes to prevent one case of brachial plexus injury

#### Case 1: Screen 100,000 pregnant women

**Assume:**
- The prevalence of gestational diabetes is 4% (average risk)
- 30% of women with gestational diabetes require insulin (assuming aggressive criteria)
- Tight control of glucose reduces the development of macrosomia (birthweight >4,000 grams) from 36% to 9% among women treated with insulin
- Infants weighing greater than 4,000 grams at birth have a 3.5% rate of shoulder dystocia injury
- There is no benefit from treating women who do not require insulin

<table>
<thead>
<tr>
<th>Number detected by screening</th>
<th>20,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis of gestational diabetes</td>
<td>4,000</td>
</tr>
<tr>
<td>Number requiring insulin</td>
<td>1,200</td>
</tr>
<tr>
<td>Macrosomic infants (treatment/no treatment)</td>
<td>108/432</td>
</tr>
<tr>
<td>Shoulder dystocia injuries in macrosomic infants (treatment/no treatment)*</td>
<td>3.8/15.1</td>
</tr>
<tr>
<td>Difference: cases avoided</td>
<td>11.3</td>
</tr>
<tr>
<td>Number needed to screen†</td>
<td>8,900</td>
</tr>
</tbody>
</table>

#### Case 2: Same as Case 1, except that we vary the prevalence of gestational diabetes to 6% (high-risk population)

<table>
<thead>
<tr>
<th>Diagnosis of gestational diabetes</th>
<th>6,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number requiring insulin</td>
<td>1,800</td>
</tr>
<tr>
<td>Macrosomic infants (treatment/no treatment)</td>
<td>152/648</td>
</tr>
<tr>
<td>Shoulder dystocia injuries in macrosomic infants (treatment/no treatment)*</td>
<td>5.7/22.7</td>
</tr>
<tr>
<td>Difference: cases avoided</td>
<td>17.0</td>
</tr>
<tr>
<td>Number needed to screen†</td>
<td>5,900</td>
</tr>
</tbody>
</table>

#### Case 3: Same as Case 2, except that we assume that we treat 50% of women with GDM with insulin

<table>
<thead>
<tr>
<th>Diagnosis of gestational diabetes</th>
<th>3,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number requiring insulin</td>
<td>3,000</td>
</tr>
<tr>
<td>Macrosomic infants (treatment/no treatment)</td>
<td>270/1,080</td>
</tr>
<tr>
<td>Shoulder dystocia injuries in macrosomic infants (treatment/no treatment)*</td>
<td>9.5/37.8</td>
</tr>
<tr>
<td>Difference: cases avoided</td>
<td>28.3†</td>
</tr>
<tr>
<td>Number needed to screen†</td>
<td>3,600†</td>
</tr>
</tbody>
</table>

* Injuries category rounded to nearest 0.1.
† All NNS calculations are rounded upward to nearest hundred.
‡ Assuming a further 10% increase in cases avoided from treatment of women with GDM but without macrosomic infants, then cases avoided for Case 3 would be 31.1 and NNS would be about 3,300 (best case scenario).
an untreated control arm, masked obstetrical care, and measurement of health (not just intermediate) outcomes. These studies should also monitor and report harms and costs associated with screening and intensive treatment. The National Institute for Child Health and Human Development is sponsoring one such study—the Maternal-Fetal Medicine Network multicenter trial of treatment of mild GDM, involving approximately 2,400 women—to be completed in 2004.

Screening for GDM is contentious. The reason for this controversy is largely a lack of high-quality research addressing the central issues. Only good research can end the controversy and inform us how to best serve our patients.

Grant Support: This study was conducted by the RTI-UNC Evidence-based Practice Center under contract to the Agency for Healthcare Research and Quality (Contract No. 290-97-0011), Rockville, MD.

References


17. Sernner M, Naylor DC, Gare DJ, et al. Impact of increasing carbohydrate intolerance on maternal-fetal outcomes in 3637 women without gestational diabetes. The Toronto Tri-Hospital Gestational


76. Olefsky JM, Reaven GM. Insulin and glucose responses to identical oral glucose tolerance tests performed forty-eight hours apart. Diabetes. 1974;23:449–453.


