Gestational diabetes screening with the new IADPSG guidelines: a cost-effectiveness analysis

John F. Mission, MD; Mika S. Ohno, MD; Yvonne W. Cheng, MD, MPH; Aaron B. Caughey, MD, PhD

OBJECTIVE: This study investigates the cost effectiveness of gestational diabetes mellitus screening using the new International Association of Diabetes in Pregnancy Study Group (IADPSG) guidelines.

STUDY DESIGN: A decision analytic model was built comparing routine screening with the 2-hour (2h) oral glucose tolerance test (OGTT) vs the 1-hour glucose challenge test. All probabilities, costs, and benefits were derived from the literature. Base case, sensitivity analyses, and a Monte Carlo simulation were performed.

RESULTS: Screening with the 2h OGTT was more expensive, more effective, and cost effective at $61,503/quality-adjusted life year. In a 1-way sensitivity analysis, the more inclusive IADPSG diagnostic approach remained cost effective as long as an additional 2.0% or more of patients were diagnosed and treated for gestational diabetes mellitus.

CONCLUSION: Screening at 24-28 weeks’ gestational age under the new IADPSG guidelines with the 2h OGTT is expensive but cost effective in improving maternal and neonatal outcomes. How the health care system will provide expanded care to this group of women will need to be examined.

Key words: cost effectiveness, decision analysis, gestational diabetes, International Association of Diabetes in Pregnancy Study Group guidelines, 2-hour glucose tolerance test


The prevalence of gestational diabetes mellitus (GDM) has increased alongside the prevalence of obesity in the United States, with an estimated 6-7% of pregnant patients carrying a diagnosis of GDM. GDM is also associated with numerous complications of pregnancy, including higher rates of preeclampsia, operative deliveries, macrosomia, shoulder dystocia, and birth injuries.

Diagnostic criteria for GDM are evolving. The Hyperglycemia and Adverse Pregnancy Outcomes (HAPO) study demonstrated that hyperglycemia at levels below those diagnostic for GDM were associated with adverse maternal and neonatal outcomes. For this reason, the International Association of Diabetes in Pregnancy Study Group (IADPSG) convened a workshop conference in 2008 where they recommended using new cutoffs for the 2-hour (2h) oral glucose tolerance test (OGTT) in GDM screening and diagnosis. The IADPSG selected cutoffs associated with an odds ratio of 1.75 times the mean for outcomes of increased body fat, large for gestational age, and cord serum C-peptide >90th centile, yielding the new recommended diagnostic criteria for GDM. In the United States, this would mean moving away from a screening test followed by a diagnostic test to a single universal test consisting of a fasting blood glucose followed by the 2h OGTT using a 75-g glucose load. Under these new criteria, patients would qualify for a diagnosis of GDM if their 2h OGTT exceeded ≥1 of the following thresholds: a fasting glucose >92 mg/dL, a 1-hour (1h) plasma glucose >180 mg/dL, or a 2h plasma glucose >153 mg/dL.

The adoption of these criteria is controversial. According to these criteria, an estimated 18% of patients would qualify for a diagnosis of GDM, potentially adding to the costs of care for many pregnant women in the United States. Prior cost-effectiveness analyses have found universal screening with the 1h glucose challenge test (GCT) to be more cost effective than the 2h OGTT. These studies, however, examined the previous World Health Organization (WHO) criteria used for GDM screening in pregnancy. In addition, 2 randomized controlled trials (RCT) published after these studies have demonstrated that treatment of GDM decreases the risk of maternal and neonatal adverse outcomes, with the more recent study demonstrating that treating mild GDM decreased the risks of fetal macrosomia, shoulder dystocia, cesarean delivery, and hypertensive disorders.

The goal of our study was to conduct a cost-effectiveness analysis to determine from a societal standpoint which routine GDM screening method was more cost effective: universal screening according to current American Congress of Obstetricians and Gynecologists (ACOG)
guidelines with the 1h GCT followed by a 3-hour OGTT or the new IADPSG guidelines for the 2h OGTT.

**Materials and Methods**

A decision-analytic model was designed using TreeAge Pro Software (2011 version; TreeAge Software, Inc., Williamstown, MA) to compare total costs and total maternal and neonatal quality-adjusted life years (QALYs) from a societal perspective for women undergoing routine GDM screening at 24-28 weeks with either the routine 1h GCT or the new IADPSG guidelines for the 2h OGTT (Figure 1). Given that current ACOG guidelines recommend universal GDM screening, the option of no screening was not investigated and only screening with the 2h OGTT and the 1h GCT were compared. As no human subjects were involved in creating this theoretical model, this study was exempt from institutional review board approval. Patients in each strategy were categorized into 3 groups: patients who would qualify for a GDM diagnosis by the 1h GCT criteria (group 1), the additional patients who would be diagnosed with GDM under the new screening guidelines (group 2), and patients with normal glycemic levels (group 3).

Maternal outcomes in the model included: preeclampsia, shoulder dystocia, cesarean vs vaginal delivery, and maternal death. Neonatal outcomes included: macrosomia (≥4000 g), permanent or transient brachial plexus injury (BPI), hypoglycemia, admission to a neonatal intensive care unit (NICU), hyperbilirubinemia, and neonatal death. All probabilities, costs, and utilities were derived from the literature. The incremental cost-effectiveness ratio (ICER), the ratio of health care dollars spent to health outcomes obtained, was measured in 2012 $US/QALY gained. In the United States, the range of what is considered cost-effective ranges from $50,000–100,000/QALY. Thus, we considered any ICER <$50,000/QALY as cost effective and any ICER between $50,000–100,000/QALY as marginally cost effective. Cost-effectiveness thresholds ≥$100,000/QALY were considered not cost effective.8

**Probabilities**

The baseline probabilities for preeclampsia, cesarean delivery, macrosomia, neonatal death, and NICU admission were derived from the literature.8,9,10,11 (Table 1). For the group of patients who would be diagnosed under the 2h OGTT using the new IADPSG guidelines, the baseline probabilities were derived directly from the data for HAPO group 5, representing the patients within the HAPO study whose glycemic values straddle the new IADPSG guidelines.3,11 The baseline probability of falling in...
TABLE 1
Probabilities, costs, and utilities used in model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Probabilities</th>
<th>Utilities</th>
<th>Costs</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDM treatment</td>
<td>.054</td>
<td>1</td>
<td>$1971</td>
<td>17</td>
</tr>
<tr>
<td>1h GCT screen</td>
<td>.124</td>
<td>1</td>
<td>$83</td>
<td>16</td>
</tr>
<tr>
<td>2h OGTT screen</td>
<td>1</td>
<td>1</td>
<td>$98</td>
<td>16</td>
</tr>
<tr>
<td>Missed GDM with 1h GCT</td>
<td>.554</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDM diagnosis with 1h GCT</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maternal outcomes</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NICU admission</td>
<td>.116</td>
<td>.116</td>
<td>.17</td>
<td>.17</td>
<td>.08</td>
</tr>
<tr>
<td>Hypoglycemia</td>
<td>1</td>
<td></td>
<td>$2669</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Permanent BPI</td>
<td>.067</td>
<td>.067</td>
<td>.067</td>
<td>.067</td>
<td>.067</td>
</tr>
<tr>
<td>Maternal death</td>
<td>1</td>
<td></td>
<td>$8595</td>
<td>20,27</td>
<td></td>
</tr>
<tr>
<td>Cesarean delivery</td>
<td>.338</td>
<td>.338</td>
<td>.228</td>
<td>.180</td>
<td>.175</td>
</tr>
<tr>
<td>Vaginal delivery</td>
<td>1</td>
<td></td>
<td>$100,000</td>
<td>10,25</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neonatal outcomes</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Macrosomia</td>
<td>.16</td>
<td>.0656</td>
<td>.15</td>
<td>.069</td>
<td>.802</td>
</tr>
<tr>
<td>Shoulder dystocia</td>
<td>.18</td>
<td>.0864</td>
<td>.18</td>
<td>.0864</td>
<td>.007</td>
</tr>
<tr>
<td>Macrosomia</td>
<td>.09</td>
<td>.0432</td>
<td>.09</td>
<td>.0432</td>
<td>.001</td>
</tr>
<tr>
<td>Shoulder dystocia</td>
<td>.0079</td>
<td>.00379</td>
<td>.0079</td>
<td>.00379</td>
<td>.00379</td>
</tr>
</tbody>
</table>

| Hypoglycemia | 1 |  | $2669 | 12,14 |
| Hyperbilirubinemia | 1 |  | $2213 | 12,14 |
| NICU admission | .116 | .116 | .17 | .17 | .08 | 1 | $16,622 | 3,10,20,21 |
| Maternal death | .062 | .062 | .062 | .062 | .062 | 0.92 | $90,874 | 10,23,27,28 |

GCT: glucose challenge test; GDM: gestational diabetes mellitus; NICU: neonatal intensive care unit; OGTT: oral glucose tolerance test; 1h: 1-hour; 2h: 2-hour.

group 1 was obtained from 2008 data on pregnancy hospital discharges in the United States, while the baseline probability for falling in group 2 was inferred from HAPO study data.13 As the presence of macrosomia affects the probability of shoulder dystocia, BPI, hypoglycemia, and hyperbilirubinemia, these probabilities were derived from a separate body of work that examined these factors independently in patients with GDM.12,14 Probabilities for maternal death were also derived from the literature.15

To estimate the effect of GDM treatment on these outcomes for patients in group 1, the baseline probabilities were multiplied by relative risks (RRs) taken both from a recent metaanalysis on the effects of GDM treatment on pregnancy outcomes as well as the RR for treating mild GDM.10,16 For patients in group 2, baseline probabilities were multiplied by the RR associated with treatment published in the 2009 RCT on treating mild GDM.10

**Utilities**

All utilities were derived from the literature and include utilities from the maternal and neonatal perspective (Table 1). The utility of a vaginal delivery was assumed to be one and the utility of a cesarean delivery was set to a baseline of 0.99 based on previously published research on women’s preferences in mode of delivery.26-28 Maternal death by definition was set to a utility of 0. The utility of a neonatal death from the maternal perspective was set at a baseline of 0.92, the published maternal utility of a miscarriage, and was applied over the maternal lifetime.27,29 The utility of a neonatal death from the neonatal perspective was by definition 0 and applied over the neonatal lifetime. The neonatal utility of a transient BPI was set to a baseline of 0.99 based on “brachial plexus injury that resolves within 2 months.”24 For a permanent BPI, a conservative value of 0.6 was used based on a published value for mild to moderate injury, compared with a utility of 0.45 for a severe injury.24 These utilities were applied over the neonatal lifetime. As there are no published utilities on short-term neonatal consequences such as hypoglycemia, NICU admission, or hyperbilirubinemia, these utilities were set to a conservative baseline value of 1.27 Utilities were calculated over the course of maternal life expectancy (56.1 years) and neonatal life expectancy (77.2 years) at a discount rate of 3%.9 All utilities were subject to sensitivity analysis.

**Analysis**

Baseline analysis compared routine GDM screening with the 1h GCT vs the 2h OGTT to estimate differences in maternal and neonatal outcomes, total costs, and QALYs for each strategy as well as the ICER.

Sensitivity analysis is a statistical tool that helps clarify key variables that most influence outcomes in the model and test the robustness of the model. A threshold value marks the point at which change in a variable would alter the conclusion. We tested the robustness of our model by performing univariate sensitivity analysis on all inputs and also introduced an additional variable in the sensitivity analysis to represent the percent efficacy of treatment, attempting to account for practice variations that can result in a range of outcomes with treatment.

Monte Carlo analysis is a tool that can incorporate uncertainty into the baseline stochastic model by sampling the underlying distributions for each input. One simulation represents a woman undergoing routine GDM screening with either the 1h GCT or the 2h OGTT. Her probabilities and costs are randomly chosen from a prespecified distribution determined from the literature, and this simulation is repeated with a different set of randomly chosen values, with the aggregate representing a theoretical cohort of random women.27 Based on these simulations, an acceptability curve can be constructed showing the probability of achieving cost effectiveness. To adopt the deterministic model to a stochastic one for the Monte Carlo simulation, costs were modeled using a gamma distribution, and probabilities were modeled using a beta distribution.

**Results**

Maternal and neonatal outcomes associated with GDM screening for a theoretical cohort of 100,000 women were estimated (Table 2). Screening with the 2h OGTT decreased all maternal outcomes, including preeclampsia, cesarean delivery, and shoulder dystocia. In addition, using the 2h OGTT decreased the neonatal outcomes of macrosomia, hypoglycemia, hyperbilirubinemia, and transient BPI. Using the 2h OGTT did not have an effect on permanent BPI.

The baseline cost-effectiveness analysis results show that screening with the 2h OGTT is more expensive at $12,201 vs $12,078 with 1h GCT screening but more effective shown by higher QALYs at 56.713799 vs 56.711790 with the 1h GCT (Table 3). The incremental cost/QALY is cost effective (below the cost-effectiveness threshold of $100,000/QALY) at $61,503/QALY.

**Sensitivity analysis**

Univariate sensitivity analyses were conducted on all probabilities, costs, and utilities. The model remained robust
Examining the additional percentage of pregnant women who would be diagnosed with GDM under the IADPSG guidelines (group 2) revealed that screening with the 2h OGTT was cost effective (<$100,000/QALY) as long as an additional 2.04% of screened patients would be newly diagnosed with GDM under the new guidelines (Figure 2, A).

Examining costs, the only costs that made GDM screening with the 2h OGTT no longer cost effective were the incremental cost of GDM treatment and the cost of the 2h OGTT (Figure 2, B and C). Screening with the 2h OGTT was cost effective (<$100,000/QALY) as long as the cost to treat GDM was <$2630 (Figure 2, B) and as long as the cost of the 2h OGTT was <$175.74 (Figure 2, C). To examine the interaction of the costs of the 1h and 2h screening tests, 2-way sensitivity analyses comparing these 2 variables were performed and showed that the model remained cost effective as long as the incremental cost of 2h screening over 1h screening was <$92.56 (Figure 3, A).

Univariate sensitivity analysis on the efficacy of treatment showed that at the baseline cost to treat GDM of $1971, screening with the 2h OGTT was cost effective as long as GDM treatment was at least 74.9% efficacious (Figure 2, D). Since the efficacy of treatment also affected the cost-effectiveness threshold (the lower the efficacy, the lower the cost threshold were it was no longer cost effective to screen with the 2h OGTT), 2-way sensitivity analysis was performed on treatment efficacy vs incremental cost to treat GDM (Figure 3, B). At 100% treatment efficacy, the cost to treat could be as high as $2630 and still be cost effective; this represents the baseline cost-effectiveness threshold. As the treatment efficacy decreased, the upper limit on the cost to treat GDM that was still cost effective decreased linearly.

Sensitivity analysis was performed to further examine the outcomes that had the greatest impact on the cost effectiveness of screening with the 2h OGTT. The reduction in 2 outcomes showed an effect on the cost effectiveness of screening with the 2h OGTT: the reduction in rates of preeclampsia and cesarean section in group 2 from GDM diagnosis and from subsequent treatment under the new 2h OGTT compared to no diagnosis or treatment with the 1h OGTT. Two-way sensitivity analysis was performed on the probabilities of preeclampsia in group 2 with GDM diagnosis and treatment under the 2h OGTT vs 1h OGTT screening without GDM diagnosis or treatment. The model remained cost effective as long as the incremental reduction in the preeclampsia rate was >0.55% with 2h OGTT GDM diagnosis and subsequent treatment (Figure 3, C). Likewise, using 2-way sensitivity analysis to compare the rates of cesarean in group 2 with and without treatment under the 2h OGTT vs the 1h GCT, the model was cost effective as long as the incremental reduction in the cesarean section rate was >2.7% with 2h OGTT screening, GDM diagnosis, and subsequent GDM treatment in this group (Figure 3, D).

A Monte Carlo simulation analysis was performed to simulate the outcome of 10,000 random pregnant women undergoing GDM screening. Based on these simulations, a willingness-to-pay acceptability curve showed that a cost-effectiveness threshold of $100,000/QALY yielded a 75.7% probability that screening with the 2h OGTT would be cost effective (Figure 4).

**Comment**

Our decision analytic model showed that screening with the 2h glucose tolerance test using the new IADPSG guidelines was cost effective as long as the cost to treat GDM was <$2630 or when the efficacy of treatment met at least 74.9% of...
The aforementioned cost analysis also differed in the RR reduction values used for the outcomes of cesarean section and preeclampsia. Their analyses assumed that the benefits of GDM treatment would translate to the additional patients who would newly qualify for a GDM diagnosis under the new IADPSG criteria. To calculate the effects of treatment on cesarean section and other outcomes, this group used the results from a recent metaanalysis on perinatal outcomes after GDM treatment to estimate the effects of treatment on rates of cesarean section, which found no difference in the rates of cesarean section in patients with GDM. Our analysis extrapolated the effect of treatment on cesarean section by using the results of the RCT examining treatment of mild GDM, which found a RR reduction of 0.79 in cesarean section rate with GDM treatment, and applying these results to patients in group 2 of our analysis. As indicated in our sensitivity analyses, an incremental reduction in the cesarean section rate by 2.7% in group 2 is required for our model to be cost effective, and if there were no difference in cesarean rates with GDM treatment, screening with the 2h OGTT would not be cost effective under our model.

Examining preeclampsia reduction with treatment, the prior study also derived a value 0.65 for the RR of preeclampsia with GDM treatment using data from a Cochrane review on pre-
eclampsia rates after GDM treatment.\textsuperscript{32} This value is more conservative than the RR of 0.46, which was also derived from the study investigating treating mild GDM and applied to group 2 of our analysis.\textsuperscript{8} According to our sensitivity analysis on preeclampsia rates, however, only an absolute risk reduction in preeclampsia rates by 0.55\% or 0.0055 is required to maintain cost effectiveness in our model. With a baseline preeclampsia risk of 0.0679, this would correspond to an RR of 0.92 required for our model to still be cost effective. Thus, using the more conservative RR from their analysis would still result in the 2h OGTT screening being cost effective over the 1h glucose tolerance test.

Interestingly, when broadening their analyses to include an intervention to decrease the likelihood of progression to type 2 diabetes mellitus (T2DM), screening with the 2h OGTT became cost effective using their model.\textsuperscript{31} While one RCT has shown benefit from an intervention to delay the onset of T2DM,\textsuperscript{33} rates of postpartum glucose follow-up in patients with GDM is poor,\textsuperscript{34} even despite an intervention involving intensive patient counseling to improve postpartum glucose follow-up in patients with GDM.\textsuperscript{35} Our model did not account for such a benefit in reducing the progression of GDM to T2DM, but if this finding holds true, then the benefit gained from preventing future diagnoses of T2DM could potentially further increase the cost effectiveness of our model.

Several limitations limit the interpretations of our findings. First, decision analysis has its inherent limitations in simulating real-world outcomes. In our model, some neonatal outcomes are modeled as discrete outcomes that cannot coexist. For example, a neonate in our model cannot have both hypoglycemia and hyperbilirubinemia even though both outcomes can occur in real-world situations. Adding permutations...
for all possible combinations of outcomes studied would make decision analytic models too unwieldy for practical use. Our study did, however, account for the association of variables that are very closely related (i.e., macrosomia, shoulder dystocia, and BPI). In addition, we were unable to model the potential effects of maternal GDM on long-term outcomes such as future obesity and T2DM risk as previously discussed. If such findings were incorporated into our model, the findings would be even more robust.

The findings of our study also depend on the effectiveness of GDM treatment in reducing rates of maternal and neonatal complications. The cost effectiveness of our model was particularly sensitive to reducing cesarean section rates and pre-eclampsia rates in patients who would be incrementally diagnosed with GDM under the 2h OGTT as previously described in our sensitivity analysis. Although we inferred the benefits of treatment from a RCT investigating the benefits of GDM treatment on patients with mild GDM so as to more accurately reflect a population with less severe hyperglycemia, no current data exist that examine the benefits of GDM treatment on maternal and neonatal outcomes for patients who would be incrementally diagnosed with GDM under the new IADPSG guidelines.

Another limitation involves our inability to model certain costs in our analysis. Antenatal admissions can add to the overall cost of managing women with GDM but were not included in this analysis as the effects of GDM treatment on antenatal hospitalization rates are unclear. One cost-consequence analysis of treating patients with “glucose intolerance of pregnancy” by pre-1998 WHO criteria found no difference in antenatal inpatient admission rates. Given that glycemic levels in patients may be less severe in patients who would be incrementally diagnosed under the new screening criteria, it is unclear if rates of antenatal admission would be higher with GDM treatment. Furthermore, while the costs of screening accounted for the opportunity cost of lost time from work, we were unable to model the costs of increased antenatal surveillance for patients. Finally, the values used for costs of GDM treatment in our study come from a review published in 2000 as more recent data on the costs of GDM treatment in the United States were lacking in the literature. As our model was sensitive to the costs of GDM treatment, more recent data would be useful in assuring the cost effectiveness of screening with the IADPSG guidelines.

Although our study found that universal screening with the newly proposed IADPSG guidelines for the 2h OGTT is more costly, our study contributes to a body of literature that suggests that this GDM screening method may be cost effective under certain circumstances. Our model is robust over a wide range of inputs for the parameters we investigated with several notable exceptions that warrant further study. As the cost of treatment is one of the inputs to which our model is sensitive, research into methods to decrease costs in management, supplies, and treatments would be an important step in increasing the cost effectiveness of GDM screening. The cost effectiveness of our model is also dependent on a reduction in both preeclampsia and cesarean section rates resultant from a new GDM diagnosis with the 2h OGTT and subsequent treatment. Further studies to demonstrate the benefit of treatment in patients who would be diagnosed with newly suggested GDM criteria are needed. Furthermore, the emerging field of developmental origins of human disease suggests that glucose intolerance during pregnancy may be associated with long-term outcomes such as T2DM, hypertension, and obesity. Long-term studies to investigate the potential link between GDM and adult outcomes for the children of patients with GDM are needed. Finally, our model demonstrates that only an additional 2% of patients need to be incrementally diagnosed with GDM for this screening to still remain cost effective. This is much lower than the anticipated total 18% of patients who would be diagnosed with GDM under the new guidelines, and as the incidence of obesity swells in the United States, the additional percentage of patients diagnosed under the IADPSG guidelines will continue to increase.

REFERENCES


OCTOBER 2012 American Journal of Obstetrics & Gynecology 326.e8
23. Phibbs CS, Schmitt SK. Estimates of the cost and length of stay changes that can be attributed to one-week increases in gestational age for premature infants. Early Hum Dev 2006;82:85-95.