Impact of gestational diabetes mellitus screening strategies on perinatal outcomes: A simulation study

Maicon Falavigna a,*, Isaias Prestes a, Maria I. Schmidt a, Bruce B. Duncan a, Stephen Colagiuri b, Gojka Roglic c

a Post Graduate Program in Epidemiology, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
b The George Institute for Global Health, University of Sydney, Sydney, Australia
c Chronic Diseases and Health Promotion, World Health Organization, Geneva, Switzerland

A R T I C L E   I N F O

Article history:
Received 14 September 2012
Received in revised form
26 November 2012
Accepted 13 December 2012
Published on line 15 January 2013

Keywords:
Gestational diabetes mellitus
Screening
IADPSG
Evidence-based medicine
Simulation

A B S T R A C T

Aims: To evaluate the impact on perinatal outcomes of universal gestational diabetes (GDM) screening based on 1999 WHO and IADPSG diagnostic criteria; to assess the quality of the evidence (GRADE) to support GDM screening.

Methods: Simulation of a hypothetical cohort of community-based pregnant women with 10% GDM prevalence (1999 WHO). Most parameters were obtained from recent systematic reviews.

Results: Compared to no screening, screening based on 1999 WHO criteria (followed by treatment) reduced the incidence of large for gestational age (LGA) neonates by 0.53% (95% CI 0.37–0.74%; NNS = 189) and of preeclampsia by 0.27% (0.10–0.45%; NNS = 376). Screening based on IADPSG criteria reduced incidences by 0.85% (0.54–1.29%; NNS = 117) and by 0.39% (0.15–0.65%; NNS = 257), respectively. Compared to screening based on 1999 WHO criteria, screening with IADPSG criteria reduced the incidence of LGA by 0.32% (0.09–0.63%; NNS = 309) and of preeclampsia by 0.12% (0.01–0.25; NNS = 808). The quality of evidence for both screening approaches is very low.

Conclusions: Universal screening for GDM has only a modest impact on pregnancy outcomes. The impact of screening based on IADPSG (vs. WHO, 1999) criteria is slightly larger. However, costs and resources should also be considered in local selection of a screening approach.

© 2012 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Gestational diabetes mellitus (GDM) has been defined as glucose intolerance of variable severity with onset or first recognition during pregnancy [1,2]. Although this definition has been largely accepted, the precise level of glucose intolerance characterizing GDM has been controversial over the last three decades.

Currently, two main diagnostic criteria for GDM are those, which have been recommended by the World Health Organization (1999) (WHO) and those recently proposed by the International Association of Diabetes in Pregnancy Study Groups (IADPSG). Both diagnostic criteria are based on a 2 h
75 gOGTT. In their 1999 Report, the WHO reiterated their criteria of classifying GDM with a 2 h plasma glucose ≥7.8 mmol/l (or 140 mg/dl) [3]. The IADPSG criteria, derived from the Hyperglycemia and Adverse Pregnancy Outcomes (HAPO) study [4], classify as GDM women with elevated values at any of the following three moments: a fasting glucose ≥5.1 mmol/L (92 mg/dl), a one hour result of ≥10.0 mmol/L (180 mg/dl), or a two hour result of ≥8.5 mmol/L (153 mg/dl). Both criteria predict important adverse outcomes, such as large for gestational age (LGA) neonates and preeclampsia [4,5].

Although moderate to high quality evidence supports treatment of GDM [6–8], prospective studies comparing outcomes in women screened versus not screened for GDM have not been undertaken. Nonetheless, screening asymptomatic pregnant women for GDM is a standard procedure in most parts of the world. The increasing prevalence of GDM, which probably results from the obesity epidemic [9]; the documented increase of risk of adverse pregnancy outcomes with GDM [5]; and the effective reduction of this risk with treatment [8] have stimulated medical associations to promote screening programs. In so doing, the IADPSG criteria, which define a milder and much more prevalent hyperglycaemia, are gaining increasing acceptance [10].

Within this scenario, evaluation of the impact of detecting and treating GDM is needed. In the absence of prospective controlled evaluations, one possibility is to model the impact of screening strategies by simulating, with the best possible existing data, the outcomes of a hypothetical cohort of women, when screened and not screened.

Our aim is thus to evaluate the impact of universal screening based on the 1999 WHO and IADPSG diagnostic criteria in a simulation cohort, combining data from observational and experimental studies available in the literature. As part of this process, we also aim to assess the quality of the evidence for these screening approaches according to the grading of recommendations assessment, development and evaluation (GRADE) working group guidelines [11,12].

2. Methods

This study is part of the support material prepared for the WHO consultation on the diagnosis and screening of gestational diabetes mellitus. It is based on two systematic reviews, one which evaluated the association of GDM (as diagnosed by the 1999 WHO and by the IADPSG criteria) with adverse pregnancy outcomes [5] and the other which assessed the effectiveness of GDM treatment [8]. We created a model to simulate the experience of a cohort of pregnant women undergoing universal screening in order to assess the impact of these two diagnostic criteria. Additionally, we assessed the quality of the evidence for these two screening approaches according to the GRADE working group guidelines [11,12].

2.1. Model description

2.1.1. Population

Our hypothetical population consisted of all pregnant women attending general prenatal care services who lacked a previous diagnosis of diabetes mellitus outside of pregnancy.

2.1.2. Interventions (screening approaches)

The two universal screening approaches, evaluated against no screening, were:

- that based on the 1999 WHO criteria;
- that based on the IADPSG criteria.

Both assumed subsequent treatment of those detected. This treatment, as considered in the model, was by definition that employed in clinical trials producing the results summarized in the systematic review [8]. It thus involved lifestyle modifications, pharmacological intervention when deemed necessary, glucose monitoring and more intensive obstetric care.

2.1.3. Outcomes of interest

Robust data do not exist for most perinatal outcomes related to GDM. Thus, important outcomes such as maternal and perinatal mortality, birth trauma and shoulder dystocia were not evaluated. Although associations for macrosomia – an outcome commonly assessed in randomized trials of GDM – were available with respect to the 1999 WHO criteria, they were not included as data were sparse for the IADPSG criteria [5]. Adequate data were available for LGA neonates, preeclampsia and caesarean section, and thus simulations were performed only for these outcomes.

2.1.4. Model parameters

Parameters used in model simulations, including their most plausible values and the upper and lower limits used to estimate their population distributions, are presented in Table 1.

As the frequency of GDM varies worldwide, we considered a prevalence of 10% for the 1999 WHO criteria, which is close to that found in the HAPO study [4] and lies well within the range of published estimates (5–15%) [5]. Less data are available to estimate the prevalence according to the IADPSG criteria. In large cohort studies, this prevalence was 25–130% greater [13,14] than that of the 1999 WHO criteria. We used a value of 50%, consistent with the findings of the HAPO study, for this model [4,15].

We estimated that compliance would be 10% lower in this simulation of a real world setting than in the context of the clinical trials included in the systematic review.

We defined baseline risks of the three outcomes in women without GDM, using data obtained from a systematic review for the 1999 WHO criteria [5]. Rates expected for having an LGA neonate, preeclampsia and caesarean section in women without GDM were estimated at 9%, 4.5% and 18.5% respectively [5]. We calculated baseline risk for women with GDM according to the 1999 WHO criteria by combining these rates with relative risks (Table 1) for these outcomes obtained from the same source [5].

To make results comparable across simulations of the different screening strategies, we fixed the baseline (without treatment) risk of the outcome being evaluated in the cohort to be equal across simulations, applying the following equation:

\[
(1 - P_{\text{WHO}}) \times R_{\text{WHO-negative}} + P_{\text{WHO}} \times R_{\text{WHO-positive}} = (1 - P_{\text{IADPSG}}) \times R_{\text{IADPSG-negative}} + P_{\text{IADPSG}} \times R_{\text{IADPSG-positive}}
\]
where $P_{\text{WHO}}$: prevalence of GDM according to the 1999 WHO criteria; $P_{\text{IADPSG}}$: prevalence of GDM according to the IADPSG criteria; $R_{\text{WHO\_negative}}$: baseline risk of women without GDM based on the 1999 WHO criteria; $R_{\text{WHO\_positive}}$: baseline risk of women with GDM based on the 1999 WHO criteria; $R_{\text{IADPSG\_negative}}$: baseline risk of women without GDM based on the IADPSG criteria; $R_{\text{IADPSG\_positive}}$: baseline risk of women with GDM based on the IADPSG criteria.

To estimate baseline risks for screen negative and positive women according to the IADPSG strategy simulation, we then applied the following equation, which uses baseline risks for the 1999 WHO strategy. Considering that:

$$P_{\text{IADPSG}} = P_{\text{WHO}} \times PR_{\text{IADPSG\_WHO}}$$

$$R_{\text{IADPSG\_positive}} = R_{\text{IADPSG\_negative}} \times RR_{\text{IADPSG}}$$

$$R_{\text{WHO\_positive}} = R_{\text{WHO\_negative}} \times RR_{\text{WHO}}$$

then

$$R_{\text{IADPSG\_negative}} = \frac{(P_{\text{WHO}} \times R_{\text{WHO\_negative}} \times RR_{\text{WHO}}) + (1 - P_{\text{WHO}}) \times R_{\text{WHO\_negative}}}{(1 - P_{\text{WHO}} \times PR_{\text{IADPSG\_WHO}}) + (P_{\text{WHO}} \times PR_{\text{IADPSG\_WHO}} \times RR_{\text{IADPSG}})}$$

where $PR_{\text{IADPSG\_WHO}}$ = ratio for prevalence increase of GDM with IADPSG criteria compared to 1999 WHO Criteria; $RR_{\text{IADPSG}}$ = relative risk for the IADPSG criteria; $RR_{\text{WHO}}$ = relative risk for the 1999 WHO criteria.

These estimated baseline risks are presented in Table 1 and are close to the incidences of these outcomes observed in untreated women evaluated in the same systematic review [5].

GDM treatment benefits (Table 1) expressed as lower relative risks, were estimated using data from a separate systematic review [8].

## Table 1 – Parameters used in the main model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Base case</th>
<th>Lower limit</th>
<th>Upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDM prevalence according to 1999 WHO criteria</td>
<td>10%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GDM prevalence according to IADPSG criteria (1999 WHO × 1.5)</td>
<td>15%</td>
<td>13.0%</td>
<td>17.3%</td>
</tr>
<tr>
<td>Probability of woman with GDM receiving treatment</td>
<td>90%</td>
<td>80%</td>
<td>97%</td>
</tr>
<tr>
<td>Baseline (without treatment) risk of a given outcome in 1999 WHO criteria negative women</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGA neonate</td>
<td>9%</td>
<td>8.5%</td>
<td>9.5%</td>
</tr>
<tr>
<td>Preeclampsia</td>
<td>4.5%</td>
<td>2.9%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Caesarean section</td>
<td>18.5%</td>
<td>10%</td>
<td>29%</td>
</tr>
<tr>
<td>Relative risk of outcome for women meeting 1999 WHO criteria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGA neonate</td>
<td>1.53</td>
<td>1.39</td>
<td>1.69</td>
</tr>
<tr>
<td>Preeclampsia</td>
<td>1.69</td>
<td>1.31</td>
<td>2.18</td>
</tr>
<tr>
<td>Caesarean section</td>
<td>1.37</td>
<td>1.24</td>
<td>1.51</td>
</tr>
<tr>
<td>Baseline (without treatment) risk of given outcome in IADPSG criteria negative women</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGA neonate</td>
<td>8.75%</td>
<td>8.18%</td>
<td>9.31%</td>
</tr>
<tr>
<td>Preeclampsia</td>
<td>4.42%</td>
<td>2.81%</td>
<td>6.37%</td>
</tr>
<tr>
<td>Caesarean section</td>
<td>18.5%</td>
<td>10%</td>
<td>29.1%</td>
</tr>
<tr>
<td>Relative risk of outcome for women meeting IADPSG criteria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGA neonate</td>
<td>1.73</td>
<td>1.27</td>
<td>2.35</td>
</tr>
<tr>
<td>Preeclampsia</td>
<td>1.71</td>
<td>1.37</td>
<td>2.12</td>
</tr>
<tr>
<td>Caesarean section</td>
<td>1.23</td>
<td>1.01</td>
<td>1.51</td>
</tr>
<tr>
<td>GDM treatment benefit (Relative risk)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGA neonate</td>
<td>0.57</td>
<td>0.47</td>
<td>0.71</td>
</tr>
<tr>
<td>Preeclampsia</td>
<td>0.61</td>
<td>0.46</td>
<td>0.81</td>
</tr>
<tr>
<td>Caesarean section</td>
<td>0.90</td>
<td>0.78</td>
<td>1.05</td>
</tr>
</tbody>
</table>


a See text for calculations; limits estimated by simulation.

### 2.2. Sensitivity analysis

To check the robustness of the results, three sensitivity analyses were performed, evaluating different settings:

1. As GDM prevalence varies worldwide, we evaluated the impact of screening strategies based on different prevalence estimates. Thus, alternative models, with prevalence ranging from 5 to 15% according to the 1999 WHO criteria, were performed.

2. Due to the uncertainty regarding the increase in GDM prevalence with the IADPSG criteria, we performed alternative models assuming increases of 25%, 75% and 100% when compared to the 1999 WHO criteria.

3. Given the large size and multi-country, multi-ethnic, population-based nature of the HAPO cohort, we additionally evaluated results from the model starting with the prevalence and baseline risk found in the HAPO cohort population [4].

### 2.3. Statistical analysis

We performed Bayesian Monte-Carlo simulations.[16] For each outcome and strategy, the model performed 1,000,000 simulations. We used beta distributions for proportions; and risk and prevalence ratios were converted to natural logarithms and modeled assuming normal distributions. Detailed information about parameters values, upper and lower limits, and distributions are available in Supplementary Tables 1 and 2.

Results are presented as absolute risk (incidence) reduction and number needed to screen (NNS), with 95% credible
Table 2 – Impact (absolute risk reduction and number needed to screen) of screening strategies, compared to no screening, for having a large for gestational age (LGA) neonate, preeclampsia and caesarean section.

<table>
<thead>
<tr>
<th>Main model</th>
<th>No screening</th>
<th>1999 WHO criteria based screening</th>
<th>IADPSG criteria based screening</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Incidence (%)</td>
<td>ARR (%)</td>
<td>NNS</td>
</tr>
<tr>
<td>LGA neonate</td>
<td>(95% CI)</td>
<td>(95% CI)</td>
<td>(95% CI)</td>
</tr>
<tr>
<td>LGA neonate</td>
<td>9.48% (8.98–9.98%)</td>
<td>8.95% (8.43–9.41%)</td>
<td>0.53% (0.37–0.74%)</td>
</tr>
<tr>
<td>Preeclampsia</td>
<td>4.81% (2.96–6.81%)</td>
<td>4.54% (2.79–6.44%)</td>
<td>0.27% (0.10–0.45%)</td>
</tr>
<tr>
<td>Caesarean section</td>
<td>19.18% (9.83–29.15%)</td>
<td>18.93% (9.74–28.85%)</td>
<td>0.25% (−0.12 to 0.60%)</td>
</tr>
<tr>
<td>Model applied to the HAPO setting</td>
<td>LGA neonate</td>
<td>9.57% (8.74–9.94%)</td>
<td>0.60% (0.43–0.83%)</td>
</tr>
<tr>
<td>Preeclampsia</td>
<td>5.22% (4.79–5.06%)</td>
<td>0.30% (0.16–0.43%)</td>
<td>0.85% (0.49–4.95%)</td>
</tr>
<tr>
<td>Caesarean section</td>
<td>18% (17.4–18.11%)</td>
<td>0.26% (−0.11 to 0.60%)</td>
<td>17.63% (17.15–18.15%)</td>
</tr>
</tbody>
</table>


3. Results

The base case results are presented in Table 2. The expected rates of events when no screening is done are 9.48% (95% CI 8.98–9.98%) for having an LGA neonate and 4.81% (95% CI 2.96–6.81%) for preeclampsia.

When compared to these rates, universal screening based on the 1999 WHO criteria reduced the incidence of LGA neonates by 0.53% (95% CI 0.37–0.74%; p < 0.001) and in preeclampsia by 0.27% (0.10–0.45%; p < 0.001). The corresponding NNSs to prevent one outcome were 189 and 376, respectively. Similar comparisons with regard to screening based on the IADPSG criteria showed absolute risk reductions of 0.85% (0.54–1.29%; p < 0.001) and 0.39% (0.15–0.65%; p < 0.001), respectively, the corresponding NNSs to prevent one outcome being 117 and 257. Since treatment of GDM was not found to reduce caesarean section rates significantly in the randomized clinical trials reviewed, results obtained from simulations, in consequence, indicate no effect of screening for this outcome.

Table 3 presents information on the quality of the evidence according to GRADE. We classified the quality of the evidence as very low for all outcomes, especially due to indirect evidence, as the impact of screening was assessed only by simulation. Evidence is less strong for the IADPSG criteria because of high heterogeneity in its associations with outcomes [5]. It is stronger for the outcome of having an LGA neonate, due to the generally larger number of events observed in trials evaluating treatment and cohort studies used to evaluate screening [5,8].

When compared, treatment based on the IADPSG criteria produced a greater reduction in incidence than that based on the 1999 WHO criteria in 99.97% of the simulations done for LGA neonates, in 99.93% of those for preeclampsia and in 91.07% of those for caesarean section. The adoption of the IADPSG criteria instead of the 1999 WHO criteria would reduce the incidence of LGA neonates by 0.32% (0.09–0.63%; p < 0.001) and of preeclampsia by 0.12% (0.01–0.25; p = 0.007). However, given the small difference in incidence reduction, the NNSs to obtain these additional benefits are large, 309 and 808 for LGA and preeclampsia, respectively. The quality of the evidence for the IADPSG criteria screening strategy being superior to that of the 1999 WHO criteria is also very low (Table 4).

Sensitivity analyses applying the same model to the prevalence and baseline risk found in the HAPO study setting showed similar absolute risk reductions and NNSs for all comparisons (Table 2).

Additional sensitivity analyses were done by altering the prevalence of GDM found with the 1999 WHO criteria. As seen in Fig. 1 and in Supplementary Table 3, the NNS decreases as the prevalence of GDM increases. In settings having a low GDM prevalence (for example, 5% according to the 1999 WHO criteria), reductions in incidence are small and NNSs are large for both criteria: for the 1999 WHO criteria, 0.26% (NNS = 378) for LGA neonates and 0.13% (NNS = 753) for preeclampsia; for the IADPSG criteria, 0.44% (NNS = 229) and 0.20% (NNS = 505), respectively. However, in settings of high prevalence (15% by 1999 WHO criteria, 0.25% (NNS = 80) and 0.57% (NNS = 174).

Finally, Table 3 and Supplementary Table 3 also show that the impact of the IADPSG criteria is greater when the increase
Table 3 – GRADE evaluation of screening for gestational diabetes (GDM) based on the universal application of the 1999 WHO and the International Association of Diabetes in Pregnancy Study Group (IADPSG) criteria Screening for GDM according to the 1999 WHO and to the IADPSG criteria compared to no screening strategy in pregnancy Population: pregnant women from general population intervention: 75 g oral glucose tolerance test, with specific treatment for women diagnosed with GDM according to the 1999 WHO or the IADPSG criteria Comparison: no screening strategy outcome: adverse perinatal and maternal outcomes.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>ARR (%)</th>
<th>NNS (95% CI)</th>
<th>Quality assessment</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screening strategy based on the 1999 WHO criteria versus no screening strategy</td>
<td>LGA neonate</td>
<td>0.53% (0.37–0.74%)</td>
<td>189 (134–268)</td>
<td>Cohort-based simulated population. No serious limitations, inconsistency or imprecision; very serious indirectness (cohort simulation)</td>
<td>Very low</td>
<td>Important</td>
</tr>
<tr>
<td></td>
<td>Preeclampsia</td>
<td>0.27% (0.10–0.45%)</td>
<td>376 (223–1010)</td>
<td>Cohort-based simulated population. No serious limitations or inconsistency; very serious indirectness (cohort simulation); serious imprecision (small number of outcomes for preeclampsia in randomized clinical trials)</td>
<td>Very low</td>
<td>Critical</td>
</tr>
<tr>
<td></td>
<td>Caesarean section</td>
<td>0.25% (−0.12 to 0.60%)</td>
<td>399 (165 to –848)</td>
<td>Cohort-based simulated population. Serious limitations (unblinded trials or selective blinding for control group in the evaluation of treatment efficacy); no serious inconsistency or imprecision; very serious indirectness (cohort simulation)</td>
<td>Very low</td>
<td>Critical</td>
</tr>
<tr>
<td>Screening strategy based on the IADPSG criteria versus no screening strategy</td>
<td>LGA neonate</td>
<td>0.85% (0.54–1.29%)</td>
<td>117 (77–185)</td>
<td>Cohort-based simulated population. No serious limitations or imprecision; serious inconsistency (heterogeneity in the predictive value for IADPSG criteria); very serious indirectness (cohort simulation)</td>
<td>Very low</td>
<td>Important</td>
</tr>
<tr>
<td></td>
<td>Preeclampsia</td>
<td>0.39% (0.15–0.65%)</td>
<td>257 (154–679)</td>
<td>Cohort-based simulated population. No serious limitations; serious inconsistency (heterogeneity in the predictive value for IADPSG criteria); very serious indirectness (cohort simulation); serious imprecision (small number of outcomes for preeclampsia in randomized clinical trials)</td>
<td>Very low</td>
<td>Critical</td>
</tr>
<tr>
<td></td>
<td>Caesarean section</td>
<td>0.34% (−0.16 to 0.83%)</td>
<td>296 (120 to –622)</td>
<td>Cohort-based simulated population. Serious limitations (unblinded trials or selective blinding for control group in the evaluation of treatment efficacy); serious inconsistency (heterogeneity in the predictive value for IADPSG criteria); very serious indirectness (cohort simulation); no serious imprecision</td>
<td>Very low</td>
<td>Critical</td>
</tr>
</tbody>
</table>

in the prevalence of GDM with these criteria is greater. Assuming a prevalence of 10% according to the 1999 WHO criteria, and prevalence increases of 25%, 50%, 75% and 100% with the adoption of the IADPSG criteria, we observed incidence reductions of 0.72% (NNS = 139), 0.85% (NNS = 117), 0.98% (NNS = 102) and 1.10% (NNS = 91) of having an LGA neonate; and of 0.33% (NNS = 303), 0.39% (NNS = 257), 0.45% (NNS = 224) and 0.50% (NNS = 199) for preeclampsia, respectively.

4. Discussion

This is the first study published assessing the impact of screening using the new IADPSG criteria in comparison to the 1999 WHO criteria. Our results show that screening with subsequent treatment of gestational diabetes significantly reduces the incidence of having an LGA neonate and preeclampsia. When based on the 1999 WHO criteria, incidence reduction was 0.53% (NNS = 189) for having an LGA neonate and 0.27% (NNS = 376) for preeclampsia. When based on the IADPSG criteria, incidence reduction of having an LGA neonate was 0.85% (NNS = 117) and of preeclampsia, 0.39% (NNS = 257). Given the greater number of cases detected with screening based on the IADPSG criteria, its implementation rather than the 1999 WHO criteria would reduce the incidence of LGA neonates by 0.32% (NNS = 309) and of preeclampsia by 0.12% (NNS = 808). The quality of evidence that universal screening for GDM prevents these outcomes is very low. No significant effect on caesarean section was observed with any criteria, given that evidence of moderate quality shows that GDM treatment does not produce a clinically significant reduction in this outcome [8].

Modeling outcomes using estimates of GDM prevalence and baseline risk for outcomes from the HAPO study population did not materially change this result. Strong points of our study merit mention. This simulation study uses data from recent systematic reviews, the parameters used are objective and derived from the literature, and the process of simulation is transparent, with the code published as an appendix. We performed sensitivity analysis considering settings, including that of the HAPO study population, with different prevalences of GDM. The quality of the evidence for screening was evaluated objectively according to GRADE.

Our study also has limitations. Being a simulation study, the evidence generated is indirect. Even so, it provides useful information in settings of similar baseline risk and GDM prevalence, given that randomized trials are not available. Further, the parameters available for our simulations are imperfect. First, prediction of outcomes is limited by the heterogeneous results available for IADPSG criteria [5]. Second, the increase in the prevalence of GDM using IADPSG criteria varies substantially in the literature, with some studies describing an increase of only 25% [14,22] and others an increase of more than 100% [13,21] when compared to the 1999 WHO criteria. For the main model, we assumed a base-case of 50%, near to that found in the HAPO study. This assumption may be considered conservative as we would expect a greater reduction in incidence of outcomes when...
screening and treatment are undertaken in the setting of a greater prevalence of GDM. This was, in fact, demonstrated in our sensitivity analyses. Third, prediction of treatment effects is also uncertain, especially for the IADPSG criteria, which have never been applied in randomized controlled trials of GDM treatment. Thus, we assumed treatment effects (relative risks) to be equal for both criteria.

Yet, our results are useful for policy making as they provide the objective information needed for screening program planning and implementation. Our simulation of the impact of screening strategies in settings having different GDM prevalence was performed to permit assessment of the benefit in different scenarios (Supplementary Table 3). Additionally, we provide the code of our simulation for those who want to perform estimations for their own setting (Appendix 1).

Further research is needed. We could not adequately assess the impact of screening on other important outcomes such as perinatal mortality, shoulder dystocia and intensive care unit admission because of insufficient information to reliably quantify the impact of screening; however, it is likely that GDM screening would also have a positive effect on these outcomes. Additionally, positive screeners, if advised of their high future risk of diabetes, may adopt healthy habits, the benefits of which may also be attributed to GDM screening.

Finally, we only assessed the impact of screening strategies based on universal application of the 1999 WHO and the IADPSG criteria. Other screening strategies can be considered, such as a two step approach, currently recommended by the American College of Obstetricians and Gynecologists [23], or selective screening based on risk factors [24–27].

Although not evaluated, costs are important for screening implementation. A cost-utility analysis found that screening based on the IADPSG criteria was not cost-effective unless long term maternal benefits were also considered [28]. Another recent cost-utility analysis compared this new screening strategy with universal screening according to current American Congress of Obstetricians and Gynecologist guideline (1 h glucose challenge test followed by a 3-h OGTT) finding that the screening strategy based on the IADPSG criteria may be cost-effective for high resources settings ($61,503/QALY), but probably is too costly for most countries [29].

When evaluating screening options, it is also important to consider the negative aspects related to labeling/treating as asymptomatic women. Unfortunately, little has been published in this regard with respect to GDM. However, the diagnosis of GDM may have a negative impact psychologically and in women’s perception of their own health [30–32].

In conclusion, universal screening followed by specific GDM treatment has only a modest impact on pregnancy outcomes. Although the impact based on the IADPSG criteria is slightly larger than that based on the 1999 WHO criteria, issues of cost-effectiveness and availability of resources must also be considered in decisions related to the selection of criteria for local implementation.

Conflict of interest

No potential conflicts of interest relevant to this article were reported.

Acknowledgements and funding

Financial support was received from the World Health Organization (Registration 2010/119177-0, APW 200308151).

M.F. wrote the protocol, developed the statistical model and wrote the manuscript. I.P. reviewed the statistical model and wrote the code. B.B.D., S.C. and G.R. contributed to discussion and reviewed the manuscript. M.I.S. participated in all the aspects of the project and was the overall supervisor.

Appendix A. Supplementary data

REFERENCES