Accuracy of screening tests for gestational diabetes mellitus in Southeast Asia
A systematic review of diagnostic test accuracy studies

Sattamat Lappharat, PhD, Tippawan Liabsuetrakul, MD, PhD

Abstract

Background: To investigate the accuracy of screening tests for gestational diabetes mellitus (GDM) in Southeast Asian pregnant women.

Methods: We searched PubMed (MEDLINE), Web of Science, Cochrane Library, ClinicalTrials.gov, Google Scholar, and Google for relevant articles published in English up to November 2018 using search terms related to GDM, screening tests for GDM and diagnostic performance. The studies were independently screened and selected by both authors. The methodological quality of the included studies was independently assessed by quality assessment of diagnostic accuracy studies. A hierarchical summary receiver operating characteristic (HSROC) model was created to estimate the HSROC curve. The summary sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, and diagnostic odds ratio were calculated in a meta-analysis using bivariate random-effects model.

Results: A total of 19 studies were included in which the 100 g oral glucose tolerance test (OGTT) and 75 g OGTT were the two common reference standards for diagnosis of GDM. Most points of diagnostic performance in the HSROC 50 g GCT curve compared with the 100 g OGTT reference standard were clustered in the upper left-hand quadrant. The pooled sensitivity and specificity of the 50 g GCT were 79% (95% confidence interval [CI] 64%–89%) and 74% (95% CI 59%–85%), respectively. For the 75 g OGTT reference standard, the non-fasting 2-hour plasma glucose showed quite similar sensitivity the 50 g GCT compared with the 100 g OGTT reference standard. The pooled sensitivities and specificities of the fasting plasma glucose and hemoglobin A1c were 81% (95% CI 76%–86%) and 70% (95% CI 67%–72%), and 80% (95% CI 66%–90%) and 69% (95% CI 58%–78%), respectively.

Conclusion: Our findings indicate that the 50 g GCT using the threshold of 140 mg/dL is a good screening test for identifying GDM at 24 to 28 weeks’ gestational age for both high-risk and universal screening strategies in Southeast Asian countries. The non-fasting 2-hour PG, fasting plasma glucose or hemoglobin A1c are alternative choices for screening.

Abbreviations: AUC = area under curve, DOR = diagnostic odds ratio, FPG = fasting plasma glucose, GCT = glucose challenge test, GDM = gestational diabetes mellitus, HbA1c = hemoglobin A1c, HSROC = hierarchical summary receiver operating characteristic, OGTT = oral glucose tolerance test, PG = plasma glucose, WHO = World health organization.

Keywords: diagnostic test accuracy, gestational diabetes mellitus, screening tests, sensitivity, specificity

1. Introduction

Gestational diabetes mellitus (GDM) mostly occurs in the second and third trimesters of pregnancy due to insulin resistance and glucose intolerance during pregnancy. \[1,2\] GDM has become a global public health concern due to potentially serious short- and long-term effects on both the pregnant women and their infants including pre-eclampsia, neonatal hypoglycemia, fetal growth, fetal macrosomia, and increased risk of developing future...
diabetes in both mothers and babies.\textsuperscript{1–5} The global GDM prevalences range from 1\% to 28\% depending on population characteristics, ethnicities, genetic factors, and screening and diagnostic methods or criteria used.\textsuperscript{1–5,7} Two review articles reported that Non-Caucasians, particularly Asian ethnicities, had higher rates of GDM than Caucasians.\textsuperscript{6,7}

The oral glucose tolerance test (OGTT) has been widely used as a reference standard for diagnosis of GDM, and is normally performed at a late gestational age (24–28 weeks) by either a two-step approach with a 50g glucose challenge test (GCT) followed by a 3-hour 100g OGTT or a one-step 2-hour 75g OGTT. The OGTT requires fasting for at least 8 hours before the procedure\textsuperscript{1,8} and therefore, screening tests with no requirement of fasting are preferred. The use of 50g GCT has been widely studied as an index test for screening for GDM, but previous studies have reported accuracy inconsistencies with the GCT across the world depending upon the application of the tests, cut-off thresholds, and population characteristics.\textsuperscript{1–5} The use of a 75g glucose load in a non-fasting state (non-fasting 75g 2-hour PG), following the Diabetes in Pregnancy Study Group of India criteria, has also been recently studied.\textsuperscript{9} Due to the shortcomings of glucose loading with its gastrointestinal side effects on pregnant women, the fasting plasma glucose (FPG) and hemoglobin A1c (HbA1c) tests have been alternatively used as screening tests for GDM, but their usefulness is still uncertain.\textsuperscript{10–19} Apart from maternal investigation using blood plasma, the fetal biometry measured by ultrasonography has been studied for detection of GDM.\textsuperscript{17–19}

Although the detection of GDM is crucial and GDM testing is recommended by the World Health Organization (WHO), a recommendation on whether or how to screen GDM is not definitely determined and routine screening is not suggested. The WHO suggests that identification of effective screening strategies for GDM is prioritized for research in low- and middle-income countries.\textsuperscript{20} To date, there is a lack of uniformity in screening and diagnostic methods of detecting GDM, even though screening and diagnosis of GDM is currently applied in routine clinical practice. Due to the high prevalence of GDM and its related complications in the WHO Southeast Asia Region,\textsuperscript{6,7,18,21} this systematic review aimed to investigate the accuracy of screening tests for screening GDM in Southeast Asian pregnant women.

2. Methods

This review was conducted in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis of Diagnostic Test Accuracy: The preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy statement.\textsuperscript{22} The review protocol was registered with the International Prospective Register of Systematic Reviews (CRD42018114375) and approved by the Institute Ethics Committee of the Faculty of Medicine, Prince of Songkla University (REC.61-337-18-1).

2.1. Eligibility criteria

We included various types of studies, such as cross-sectional studies, retrospective and prospective cohort studies, or randomized controlled trials, which had been conducted in countries included in the WHO Southeast Asia Region, and the studies had assessed the accuracy of screening tests for gestational diabetes mellitus. Case-control studies were excluded due to selection and performance bias.\textsuperscript{23}

We selected studies in which Southeast Asian pregnant women of any gestational age and risk of GDM, who had received screening tests for GDM during their prenatal visits. Those with known diabetes mellitus before pregnancy or having a history of GDM were excluded. Both the 2- and 1-step approaches for screening for GDM regardless of type of index test or reference standard used were considered.

2.2. Search strategy and data sources

We searched PubMed (MEDLINE), Web of Science, Cochrane Library, and ClinicalTrials.gov for relevant articles published in English up to November 2018 using search terms related to GDM, counties in Southeast Asia, the aforementioned index tests and diagnostic performance. All search term details are provided in Appendix 1 (http://links.lww.com/MD/F201) as supplementary material. We also conducted a manual search using Google Scholar and Google after retrieving articles from the database. Duplicate articles were identified and removed before assessing the remaining articles.

2.3. Study selection

Both review authors independently screened the titles and abstracts of all search results that met the eligibility criteria using Rayyan software.\textsuperscript{24} In cases where the titles or abstracts had insufficient information to either include or exclude, the full texts were retrieved and assessed independently. Disagreements and discrepancies were resolved through discussion. The number of included and excluded records was mapped with a preferred reporting items for systematic review and meta-analysis flow diagram.\textsuperscript{25}

2.3.1. Data extraction and management

An extraction form was developed with the following information: study details (title, first author, year of publication, country); study characteristics (study design, study site, sample size); participants’ characteristics (age, gestational age); index tests characteristics (gestational age, type of GDM screening, type of index test, cut-off value); reference standard test characteristics (gestational age, interval time between index test and reference standard test, glucose loading, diagnostic criteria, cut-off value); and study results (GDM prevalence, true-positive, false-positive, false-negative, true-negative). The data from the included studies were extracted independently. When data were detected to be insufficient or inconsistent to construct a 2 × 2 contingency table,\textsuperscript{25} we contacted the authors for further information. Any discrepancies were resolved by discussion and consensus.

2.3.2. Assessment of methodological quality

The two reviewers independently graded the methodological quality of the included studies, using the signaling questions of the Quality Assessment of Diagnostic Accuracy Studies 2 assessment tool for the 4 key domains (patient selection, index test(s), reference standard, and flow and timing). Each domain was assessed for the risk of bias and applicability, for which each study was classified in all domains as “low risk of bias” and “low concern” as having high methodological quality.\textsuperscript{26} Differences were resolved through discussion.

2.3.3. Statistical analysis and data synthesis

The sensitivities and specificities at multiple thresholds of an individual index test
with the same set of reference standards were plotted, and then
the optimum threshold of each index test was chosen. The data of
the selected optimum thresholds of the index tests were analyzed
and overall sensitivities and specificities of various index tests
with both reference standards were plotted by coupled forest
plots.

A hierarchical summary receiver operating characteristic
(HSROC) model was constructed to estimate a HSROC curve.[27]
The HSROC model provides equivalent summary estimates for
sensitivity and specificity and 95% confidence and prediction
regions which describe the uncertainty of the summary sensitivity
and specificity. The confidence region is related to the summary
estimates of sensitivity and specificity jointly in the HSROC space
without consideration of between-studies heterogeneity. The
prediction region refers to potential values of sensitivity and
specificity that predict the summary sensitivity and specificity of a
future study reflecting the between-studies heterogeneity.[28]

The summary sensitivity, specificity, positive likelihood ratio,
negative likelihood ratio, and diagnostic odds ratio (DOR) were
calculated in a meta-analysis using a bivariate random-effects
model.[29,30] The heterogeneity of the studies was estimated by I2
and visual inspection of forest plots.[31] A meta-regression
considering covariates, namely gestational age at screening,
country, sample size, diagnostic criteria of reference standard,
and prevalence of GDM, was performed. The possibility of
publication bias was tested by using Deek funnel plot.[32] A P-
value of <.05 was considered statistically significant for all
analyses, whereas the Deek funnel plot considered a value of
P < .10 as statistically significant. The Review Manager Version
5.3 program (Copenhagen: The Nordic Cochrane Centre, The
Cochrane Collaboration, 2014) was used to construct coupled
forest plots. Analyses were performed with Stata Version 15.1
software (StataCorp, College station, Texas, USA) using the
“midas” and “metandi” commands.

3. Results

3.1. Study selection and study characteristics

Of 286 studies found, 21 studies[14,16,33–51] met the criteria, but
the data of 2 studies[36,44] were insufficient to be extracted
resulting in a total of 19 studies[14,16,33–35,37–43,45–51] being
included in the quantitative analyses. The flow chart of literature
screening and selection process is shown in Figure 1. Two
common reference standards for diagnosis of GDM, the 3-hour
100 g OGTT and the 2-hour 75 g OGTT, were found. The characteristics of the 11 included studies[33–43] which examined
the 100 g OGTT reference standard are shown in Table 1. These
studies were conducted in Thailand, India, and Nepal. Of the 11
studies, 10[33–36,38–43] of them used the 50 g GCT test for GDM
screening at a gestational age of 24-28 weeks or less. The criteria
of the reference standard used for GDM diagnosis were either the
Carpenter-Coustan criteria or the National Diabetes Data Group
criteria. The characteristics of the 10 included studies[14,16,44–51]
using the 75 g OGTT reference standard are shown in Table 2.
Most of these studies were conducted in India using a variety of
index tests, namely the FPG, non-fasting 2-hour PG, and HbA1c
tests, and they were given at a gestational age lower than 24-28
weeks. For diagnosis of GDM the criteria of the International
Association of the Diabetes and Pregnancy Study Groups and
WHO were used.

3.2. Assessment of methodological quality of included
studies

The quality assessment of the included studies is summarized in
Figure 2. More than half were at low risk of bias and low
applicability concerns in all domains. Of the 21 studies,[14,16,33–
51] 14 studies[14,16,33–40,43,44,46,48–51] were at low risk of bias
for participant selection and 7 studies[34–36,41,42,45,47] were at
unclear risk of bias due to insufficient information of exclusion
criteria. High applicability concerns of patient selection were
found in four studies[33,34,36,41] because only women having
positive index tests were tested with a reference test. Thirteen
studies[33,37,39,41,43,45,46,48,50] were at low risk of bias for the
index test and eight studies[14,16,38,42,44,47,49,51] were at high risk
of bias due to either unclearly pre-specified thresholds used or
interpreting the results of the index test without being blinded. A
low risk of bias for the reference standard was shown in 17
studies[14,16,33–40,42–51] while the other four studies[33,34,36,41] 
were at high risk because the interpretation of the reference
standard results was done without being blinded. All studies
[14,16,33–51] were judged to have only low applicability
concerns for both index test and reference standard. Eleven
studies[14,16,35,37,39,40,42,43,46,47,51] were at low risk of bias for
the flow and timing of the study and ten studies[33,34,36,38,41,44,45,48–
50] were at high risk of bias because of an incomplete number of
participants at final analysis, an inappropriate interval between
reference standard and index test (over a week), or inconsistency
of descriptions in the Results tables and texts.

3.3. Findings of diagnostic test accuracy

Figure 3 presents the overall coupled forest plots of the different
index tests compared with the 3-hour 100 g OGTT and the 2-
hour 75 g OGTT as reference standards. The sensitivities and
specificities of the 50 GCT at the threshold of 140 mg/dL
compared with the 3-hour 100 g OGTT ranged from 36% (95% confidence
interval [CI] 11%–69%) to 100% (95% CI 88%–
100%) and 23% (95% CI 16%–30%) to 92% (95% CI 90%–
94%), respectively (Fig. 3A). The sensitivities of the non-fasting
75 g 2-hour PG with the threshold of 140 mg/dL varied from 28%
(95% CI 18%–39%) to 98% (95% CI 90%–100%) compared
with the 2-hour 75 g OGTT reference standard but specificities
were consistently high (Fig. 3B). The sensitivities and specificities of both the FPG and HbA1c were similar, with the variation of
their sensitivities better than was found in the non-fasting 75 g 2-
hour PG.

The HSROC curve comparing the 50 g GCT and 3-hour 100 g
OGTT reference standards is shown in Figure 4. Most points are
clustered in the upper left-hand quadrant. The 95% confidence
region does not overlap with the diagonal line, but the 95% prediction
region does. As there were fewer than four studies
comparing the index tests to the 2-hour 75 g OGTT reference
standard, the HSROC model could not construct for the HSROC
curve.

The pooled diagnostic performances including the DOR of two
reference standards with four index tests are shown in Table 3.
The pooled sensitivity and specificity of the 50 g GCT with 3-hour
100 g OGTT reference standard with nine studies involving 4,176
pregnant women were 79% (95% CI 64%–89%) and 74% (95% CI 59%–85%), respectively. The area under curve (AUC)
was 0.83 (95% CI 0.80–0.86) and the DOR was 10 (95% CI 5–
23), indicating high heterogeneity ($I^2$=99%). No publication
bias was found ($P=.40$). There were no statistically significant covariates revealed in the meta-regression analysis.

For the 2-hour 75g OGTT reference standard, the non-fasting 75g 2-hour PG was examined in 3 studies involving 2,767 pregnant women, and found quite similar sensitivity of the 50g GCT comparing to the 3-hour 100g OGTT reference standard. High specificity with an AUC of 0.98 (95% CI 0.96–0.99) and DOR with extremely wide confidence intervals for the non-fasting 75g 2-hour PG were found. The diagnostic performances and DORs of the FPG and HbA1c tests compared to the 2-hour 75g OGTT were similar. For the FPG, three studies involving 2,514 pregnant women showed pooled sensitivity and specificity of 81% (95% CI 76%–86%) and 70% (95% CI 67%–72%), respectively, with a DOR of 10 (95% CI 7–14) with AUC of 0.83 (95% CI 0.79–0.86). The pooled sensitivity and specificity of the HbA1c test in 2 studies involving 1,107 pregnant women were 80% (95% CI 66%–90%) and 69% (95% CI 58%–78%), respectively, with a DOR of 9 (95% CI 5–16) and AUC of 0.81 (95% CI 0.77–0.84). There was no potential publication bias for the aforementioned index tests compared with the 2-hour 75g OGTT reference standard ($P=.30$). Meta-regression could not be performed due to too few studies to conduct the analysis.

4. Discussion

Two common reference standards, the 3-hour 100g OGTT and the 2-hour 75g OGTT were used to diagnosis of GDM, and we found various index tests using the 50g GCT followed by non-fasting 75g 2-hour PG, FPG, and HbA1c in GDM screening in Southeast Asia. The majority of studies were found to have a low...
Table 1
Characteristics of included studies by 100 g OGTT reference standard.

| Study                  | Country       | Study design     | Index test   | GA (wk)   | Index test cut-off | Diagnostic criteria | GDM prevalence | No. of women | Screening group |
|------------------------|---------------|------------------|--------------|-----------|-------------------|--------------------|----------------|--------------|----------------|
| Jirapinyo 1993 [32]    | Thailand      | Prospective study| 50 g GCT     | 24–28     | 140–150 mg/dL     | NDDG               | 10.6%          | 396          | High-risk      |
| Prasila 1993 [30]      | Thailand      | Prospective study| 50 g GCT     | 24–28     | 140 mg/dL and 5.6%| NDDG               | 7.2%           | 334          | Universal      |
| Mathai 1994 [34]       | India         | NA               | 50 g GCT     | 24–28     | 130–150 mg/dL     | CC                 | 4.7%           | 232          | Universal      |
| Thitidilok 1995 [42]   | Thailand      | NA               | 50 g GCT     | 24–28     | 140–150 mg/dL     | NA                 | 7.6%           | 304          | High-risk      |
| Chanprapaph 2004 [34]  | Thailand      | Retrospective study| 50 g GCT    | <24–28    | 140 mg/dL         | NDDG               | 7.1%           | 411          | Universal      |
| Juntarit 2007 [40]     | Thailand      | Diagnostic study | 50 g GCT     | 24–28     | 130–150 mg/dL     | CC                 | 28.6%          | 598          | Universal      |
| Punthumapol 2008 [41]  | Thailand      | Retrospective study| 50 g GCT    | <24–28    | 179 mg/dL         | NDDG               | 13.2%          | 1,114         | High-risk      |
| Poornar 2013 [39]      | India         | Prospective study| 50 g GCT/ FPG| <24–28    | 130–140 mg/dl and 80–85 mg/dl | CC | 7.2% | 500 | Universal |
| Siribaddana 1998 [49]  | Sri Lanka     | Prospective study| 50 g GCT     | 24–28     | 140 mg/dL and 4.9%–5.1% | NDDG               | 24.5%          | 200          | High-risk      |
| Basnet 2018 [33]       | Nepal         | Cross-sectional  | 50 g GCT     | <24–28    | 130–140 mg/dL     | CC                 | 5.4%           | 685          | Universal      |
| Khan 2018 [37]         | India         | Non-fasting 75 g 2-h PG | 24–28 | 140 mg/dL | CC | 13.0% | 200 | Universal |

50 g GCT = 50 grams glucose challenge test, 75 g OGTT = 75 grams oral glucose tolerance test, 100 g OGTT = 100 grams oral glucose tolerance test, ADA = American Diabetes Association, CC = Carpenter-Coustan, FPG = fasting plasma glucose, GA = gestational age, HbA1c = Hemoglobin A1c, IADPSG = International Association of the Diabetes and Pregnancy Study Groups, PG = plasma glucose, NDDG = National Diabetes Data Group, NA = non-available, WHO = World Health Organization.

Table 2
Characteristics of included studies by 75 g OGTT reference standard.

| Study                  | Country      | Study design     | Index test   | GA (weeks) | Index test cut-off | Diagnostic criteria | GDM prevalence | No. of women | Screening group |
|------------------------|--------------|------------------|--------------|------------|-------------------|--------------------|----------------|--------------|----------------|
| Sinbadana 1998 [30]    | Sri Lanka    | Prospective study| 50 g GCT     | 24–28     | 140 mg/dL         | WHO 1985           | 5.5%           | 721          | Universal      |
| Senanayake 2006 [47]   | Sri Lanka    | Comparative study| FPG          | NA        | 80–126 mg/dL      | WHO 1999           | 27.7%          | 271          | High-risk      |
| Wijeyarane 2006 [31]   | Sri Lanka    | Retrospective study| FPG         | 24–28     | 80–126 mg/dL      | WHO 1999           | 16.3%          | 883          | Universal      |
| Rajput 2012 [34]       | India        | NA               | HbA1c        | 24–28     | 5.45%–5.95%       | ADA                | 7.1%           | 607          | Universal      |
| Mohan 2014 [45]        | India        | Cross-sectional  | Non-fasting 75 g 2-h PG | 24–28 | 130–150 mg/dL | WHO 1999 | 8.0% | 1,031 | Universal |
| Soumya 2015 [44]       | India        | Prospective study| HbA1c        | 24–28     | 5.3%–6.1%         | NA                 | 9.0%           | 500          | Universal      |
| Saena 2017 [32]        | India        | Cross-sectional  | Non-fasting 75 g 2-h PG | 24–28 | 140 mg/dL | WHO 1999 | 6.4% | 800 | Universal |
| Tripathi 2017 [40]     | India        | Prospective study| Non-fasting 75 g 2-h PG | 24–28 | 140 mg/dL | WHO 1999 | 6.7% | 936 | Universal |
| Agarwal 2018 [34]      | India        | NA               | FPG          | 24–28     | 76–92 mg/dl       | IADPSG             | 18.3%          | 6,520         | Universal      |
| Sharma 2018 [46]       | India        | Prospective study| FPG          | <24–28    | 84.5 mg/dl        | IADPSG             | 6.5%           | 246          | Universal      |

50 g GCT = 50 grams glucose challenge test, 75 g OGTT = 75 grams oral glucose tolerance test, 100 g OGTT = 100 grams oral glucose tolerance test, ADA = American Diabetes Association, CC = Carpenter-Coustan, FPG = fasting plasma glucose, GA = gestational age, HbA1c = Hemoglobin A1c, IADPSG = International Association of the Diabetes and Pregnancy Study Groups, PG = plasma glucose, NDDG = National Diabetes Data Group, NA = non-available, WHO = World Health Organization.

risk of bias and low applicability concerns in all domains. Our review found a wide range of sensitivities of the 50 GCT compared to the 3-hour 100 g OGTT and the non-fasting 75 g 2-hour PG compared with the 2-hour 75 g OGTT at the same threshold of 140 mg/dL. The FPG and HbA1c tests showed similar sensitivities and specificities and lower variations of sensitivities compared to the non-fasting 75 g 2-hour PG. Overall, our review indicates that the 50 g GCT using the threshold of 140 mg/dL is a good screening test for GDM at 24-28 weeks of gestation with high-risk or universal strategies. The non-fasting 75 g 2-hour PG, FPG or HbA1c tests are alternative options, but there were too few studies to come to any statistical conclusion as to their usefulness.

We found the studies focusing on the WHO Southeast Asia Region used one of the two common reference standards of the 100 g OGTT or the 75 g OGTT after fasting for the diagnosis of GDM, which earlier systematic studies also reported. Due to a lack of universal consensus regarding glucose load and diagnostic criteria for GDM, the guidelines and recommendations for screening and diagnosing GDM in pregnant women vary. A 3-hour 100 g OGTT has been proposed and used as a reference standard for diagnosis of GDM since the 1960s, which is administered by loading 100 g of oral glucose and measuring the FPG and PG levels at 1, 2, or 3 hours. The 2-hour 75 g OGTT test measures FPG and PG levels at 2 hours after loading with 75 g oral glucose. Although the 75 g OGTT test has a lower sensitivity but higher specificity, it was recommended by the WHO in 1999 as the preferred diagnostic test for GDM. This method is applied and used as a one-step test in some countries due to economical and convenient reasons.

The 50 g GCT is the most widely used screening test for GDM, used by administering a 50 g glucose load without fasting followed by a determination of PG at one hour. The common threshold of the 50 g GCT compared with the 3-hour 100 g OGTT ranges from 130 to 150 mg/dL, which is in accordance with the findings of our included studies. The best common threshold found in our systematic review was 140 mg/dL as recommended in the American Diabetes Association or WHO guidelines. We found better pooled sensitivity than specificity with the 50 g GCT test, similar to the results of previous systematic reviews, even though the criteria of the included studies and study settings in those reviews were different.
Glucose loading may cause nausea and vomiting in some pregnant women, and thus be unpleasant for them. The heterogeneity of the 50g GCT test was not resolved after meta-regression, though the known covariates were considered. This may be because meta-regression investigates the effects of multiple factors simultaneously thus nine studies may not be sufficient to reveal significant factors.

In the non-fasting 75g 2-hour PG test, PG is estimated two hours after 75g glucose loading without overnight fasting. This may cause similar side effect as the glucose loading of the 50g GCT test. In our study, we found a high variation of summary pooled sensitivity of non-fasting 75g 2-hour PG, although these results were from three studies only, and all from India. Due to the high prevalence of GDM in India reported at 16%, the use of non-fasting 75g 2-hour PG was adapted to be a national guideline of diagnostic test for screening for GDM. We found a high summary pooled specificity with a narrow confidence interval of non-fasting 75g 2-hour PG, which supports the principle of using it as a diagnostic tool. However, the study needs to be repeated with data from other countries for confirmation of clinical applications outside India.

The FPG is a plasma value which is one of abnormal findings indicating the diagnosis of GDM using for both the standard 3-hour 100g and 2-hour 75g OGTT before glucose loading. There is a consensus concerning the abnormal value that indicates a diagnosis of DM ($\geq 126$mg/dL) in general population. For pregnant women, different classifications of diagnostic criteria for GDM are recommended and various thresholds are used. Three studies conducted in Southeast Asia were found in our review which found that 84.5 to 85.0mg/dL was the same common screening threshold and gave the optimum pooled sensitivity and specificity compared to the 2-hour 75g OGTT. The thresholds of FPG for screening GDM in previous studies varied from 80 to 90mg/dL and showed a high variation of sensitivities and specificities. Compared with the same threshold of 85mg/dL, the pooled diagnostic performance of FPG in our review was lower than in a cohort study conducted in Brazil.

The HbA1c is generally used in clinical practice to diagnose and monitor DM. Owing to its properties and convenience (non-glucose loading and non-fasting), there has been substantial interest in using it as an alternative measurement for GDM screening. Our review found two studies conducted in India with thresholds of 5.45% and 5.7% which showed optimum pooled sensitivity and specificity comparable to the FPG test. A previous systematic review including eight studies from various countries showed different thresholds of HbA1c ranging from 5.4% to 6.0% with low sensitivity and high specificity for screening for GDM. Although the pooled sensitivities and specificities of both FPG and HbA1c were similar to the 50g GCT test in our review, there was evidence from only 2 or 3 included studies thus more studies using the same thresholds are required for comparisons of multiple tests to identify the suitable threshold and index test for screening GDM in the future.

The diagnostic performances of screening tests for GDM from the included studies in our review were almost all at a low risk of bias and applicability concerns. Nonetheless, there were some limitations. First, we considered high prevalence of GDM in Southeast Asia, therefore, it may be limited for generalizability. Second, a variation of thresholds was presented in each index test and we selected the optimal thresholds for our analyses which might have introduced unexpected selection bias due to our restriction process. Third, there were only a small number of studies in our meta-analyses, which mean it was difficult to perform sub-analyses to reduce heterogeneity among the studies.

Figure 2. QUADAS-2 risk of bias and applicability assessment of included studies. QUADAS-2 = quality assessment of diagnostic accuracy studies 2.
Finally, comparisons of multiple tests could not be performed again due to too few studies.

5. Clinical implications

Our study confirms that the 50 g GCT using the threshold of 140 mg/dL is the most useful screening tests for GDM in Southeast Asian pregnant women. Although the non-fasting 75 g 2-hour PG test is used widely in India, it is more commonly used as a diagnostic test rather than a screening tool. Both the FPG and HbA1c tests can be alternative methods in cases where glucose loading is not feasible. However, the number of included studies was small in our review, and more well-designed studies for diagnostic accuracy of screening tests for GDM are still required.

### Table 3

| Reference standard | Index test | Sensitivity (95% CI) | Specificity (95% CI) | LR+ (95% CI) | LR− (95% CI) | DOR (95% CI) | AUC |
|--------------------|------------|----------------------|----------------------|--------------|--------------|--------------|-----|
| 100 g OGTT         | 50 g GCT   | 0.79 (0.64, 0.89)    | 0.74 (0.59, 0.85)    | 3.00 (1.90, 4.70) | 0.29 (0.16, 0.50) | 10 (5, 23) | 0.83 |
| 75 g OGTT          | Non-fasting 75 g 2-h PG | 0.76 (0.73, 0.79) | 0.97 (0.96, 0.98) | 30.3 (13.50, 68.00) | 0.25 (0.04, 1.51) | 123 (9, 1,610) | 0.98 |
|                    | FPG        | 0.81 (0.76, 0.86)    | 0.70 (0.67, 0.72)    | 2.7 (2.40, 3.00) | 0.27 (0.21, 0.35) | 10 (7, 14) | 0.83 |
|                    | HbA1c      | 0.80 (0.66, 0.90)    | 0.69 (0.58, 0.78)    | 2.6 (2.00, 3.30) | 0.29 (0.17, 0.48) | 9 (5, 16) | 0.81 |

50g GCT = 50 grams glucose challenge test, 75g GCT = 75 grams oral glucose tolerance test, 100g OGTT = 100 grams oral glucose tolerance test, AUC = area under curve, DOR = diagnostic odds ratio, FPG = fasting plasma glucose, HbA1c = hemoglobin A1c, LR+ = positive likelihood ratio, LR− = negative likelihood ratio, PG: plasma glucose.
6. Conclusions

The 50 g GCT with the threshold of 140 mg/dL at 24 to 28 weeks of gestational age is a good screening test for identifying GDM at 24 to 28 weeks’ gestation for both high-risk and universal screening strategies in Southeast Asian countries. The non-fasting 75 g 2-hour PG test had better specificity than sensitivity, thus, it should be a diagnostic test rather than a screening test. Although both the FPG and HbA1c tests have high sensitivities and thus may be considered as alternative options for GDM screening, they still lack guidelines and threshold supports. However, all screening tests need to be confirmed by the appropriate reference standard.

Author contributions

Conceptualization: Sattamat Lappharat, Tippawan Liabsuetrakul.
Data curation: Sattamat Lappharat, Tippawan Liabsuetrakul.
Formal analysis: Sattamat Lappharat, Tippawan Liabsuetrakul.
Funding acquisition: Sattamat Lappharat, Tippawan Liabsuetrakul.
Methodology: Sattamat Lappharat, Tippawan Liabsuetrakul.
Supervision: Tippawan Liabsuetrakul.
Validation: Sattamat Lappharat, Tippawan Liabsuetrakul.
Visualization: Sattamat Lappharat, Tippawan Liabsuetrakul.
Writing – original draft: Sattamat Lappharat, Tippawan Liabsuetrakul.

References

[1] American Diabetes Association. Classification and diagnosis of diabetes. Diabetes Care 2013;36(Suppl. 1):S8–16.
[2] Juwani A, Marseille E, Lohse N, et al. Gestational diabetes mellitus: results from a survey of country prevalence and practices. J Matern Fetal Neonatal Med 2012;25:600–10.
[3] Guariguata L, Linnenkamp U, Beagley J, et al. Global estimates of the prevalence of hyperglycaemia in pregnancy. Diabetes Res Clin Pract 2014;103:176–85.
[4] Nguyen CL, Pham NM, Binns CW, et al. Prevalence of gestational diabetes mellitus in Eastern and Southeastern Asia: a systematic review and meta-analysis. J Diabetes Res 2018;2018:1–0.
[5] Tien J, McPhie AJ, Crowther CA, et al. Screening for gestational diabetes mellitus based on different risk profiles and settings for improving maternal and infant health. Cochrane Database Syst Rev 2017;8:1–54.
[6] Petry CQ. Gestational diabetes: risk factors and recent advances in its genetics and treatment. Br J Nutr 2010;104:775–87.
[7] Turno GE, Tam WH, Yang X, et al. Diabetes and pregnancy: perspectives from Asia. Diabet Med 2014;31:302–18.
[8] International Association of Diabetes Pregnancy Study Groups Consensus Panel. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care 2010;33:676–82.
[9] Phulpagar A, Deshmukh P, Gunderia A. Screening for gestational diabetes by DIPSI guidelines. Int J Biomed Res 2018;9:121–5.
[10] Sacks D. Fasting plasma glucose test at the first prenatal visit as a screen for gestational diabetes. Obstet Gynecol 2003;101:1197–203.
[11] Zhu W-W, Fan L, Yang H-X, et al. Fasting plasma glucose at 24–28 weeks to screen for gestational diabetes mellitus: new evidence from China. Diabetes Care 2013;36:2038–40.
[12] Perucchini D, Fischer U, Spinas GA, et al. Using fasting plasma glucose concentrations to screen for gestational diabetes mellitus: prospective population based study. BMJ 1999;319:812–5.
[13] Rey E, Hudon L, Michon N, et al. Fasting plasma glucose versus glucose challenge test: screening for gestational diabetes and cost effectiveness. Clin Biochem 2004;37:780–4.
[14] Souniya S, Rohilla M, Chopra S, et al. HbA1c: a useful screening test for gestational diabetes mellitus. Diabetes Technol Ther 2015;17:899–904.
[15] Fong A, Serra AE, Gabby L, et al. Use of hemoglobin A1c as an early predictor of gestational diabetes mellitus. Am J Obstet Gynecol 2014;211:e1–7.
[16] Rajput R, Yogesh Y, Rajput M, et al. Utility of HbA1c for diagnosis of gestational diabetes mellitus. Diabetes Res Clin Pract 2012;98:104–7.
[17] Perovic M, Gojnic M, Arsic B, et al. Sensitivity and specificity of ultrasonography as a screening tool for gestational diabetes mellitus. J Matern Fetal Neonatal Med 2012;25:1348–53.
[18] Perovic M, Gajic M, Arsic B, et al. Relationship between mid-trimester ultrasonar fetal liver measurement and gestational diabetes mellitus. J Diabetes 2015;7:497–505.
[19] Tantanass T, Damudilas A, Giannoulis C, et al. Sonographic assessment of fetal subcutaneous fat tissue thickness as an indicator of gestational diabetes. Eur J Obstet Gynecol Reprod Biol 2010;152:157–62.
[20] World Health Organization. WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience. Geneva: World Health Organization; 2016. 1–152.
[21] World Health Organization. Member States. SEARO. Available at: http://www.searo.who.int/about/offices/en/ [access August 1, 2019].
[22] Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLOS Med 2009;6:e1000097.
[23] Rutjes AWS, Reitsma JB, Vandenbroucke JP, et al. Case–control and two–gate designs in diagnostic accuracy studies. Clin Chem 2005;51:1335–41.
[24] Ouzzani M, Hammady H, Fentorovic Z, et al. Rayyan—a web and mobile app for systematic reviews. Syst Rev 2016;5:1–0.
[25] Macaskill P, Gatsonis C, Deeks JJ, Deeks JJ, Bossuyt PM, Gatsonis C, et al. Chapter 10: analysing and presenting results. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0 2010; The Cochrane Collaboration. Available at: http://rdtac.cochrane.org/
Whiting PF, Rutjes AS, Westwood ME, et al. Quadas-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-36.

Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med 2001;20:2865-84.

Lee J, Kim KW, Choi SH, et al. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers-part II. Statistical methods of meta-analysis. Korean J Radiol 2015;16:1188-96.

Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. J Clin Epidemiol 2006;59:1331-3.

Reitsma JB, Glas AS, Rutjes AWS, et al. Bivariate analysis of sensitivity and specificity studies, incorporating statistical dependence. BMJ 2005;331:721.

Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.

Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 2005;58:842-93.

Basnet T, Pradhan N, Koirala P, et al. Evaluation of glucose challenge test using cut off values 130 mg/dl and 140 mg/dl for gestational diabetes mellitus screening. Int J Reprod Contracept Obstet Gynecol 2018;7:801-5.

Chanprapaph P, Sutjarit C. Prevalence of gestational diabetes mellitus in women screened by glucose challenge test (GCT) at Maharaj Nakorn Chiang Mai Hospital. J Med Assoc Thai 2004;87:1141-6.

Jirapinyo M, Puavilai G, Chanprasertyotin S, et al. Predictive value of 1 hour 50 g oral glucose load screening test for gestational diabetes mellitus compared to 3 hour oral glucose tolerance test in high risk pregnant women. Asia Oceania J Obstet Gynecol 1993;19:7-12.

Juntarar W, Rueangchaimok W. 30 grams glucose challenge test for screening of gestational diabetes mellitus in high risk pregnancy. J Med Assoc Thai 2007;90:617-23.

Khan S, Bal H, Khan ID, et al. Evaluation of the diabetes in pregnancy study group of India criteria and Carpenter-Coustan criteria in the diagnosis of gestational diabetes mellitus. J Turk Soc Obstet Gynecol 2018;15:7-9.

Mathai M, Thomas TJ, Kuruvila S, et al. Random plasma glucose and the glucose challenge test in pregnancy. Natl Med J India 1994;7:160-2.

Poomalar GK, Ragaswamy V. A comparison of fasting plasma glucose and glucose challenge test for screening of gestational diabetes mellitus. J Med Assoc Thai 2014;97:549-53.

Puavilai G, Chanprasertyotin S, Jirapinyo M. An evaluation of glycosylate hemoglobin measurement by a colorimetric method as a screening test for gestational diabetes mellitus. J Med Assoc Thai 2013;95:449-50.

Punthamapol C, Tekasakul P. 30 grams glucose challenge test for screening of gestational diabetes mellitus in each trimester in potential diabetic pregnancy. J Med Assoc Thai 2008;91:787-93.

Thhatdiklo W, Techatrasak K. The screening for gestational diabetes employing 50g glucose test. J Med Assoc Thai 1995;78:526-31.

Wuthihenjarasamee K, Sirnil S, Sririppatananukul M, Hb Alc versus 50 grams glucose screening test for screening gestational diabetes mellitus. Thai J Obstet Gynecol 2014;42:22-8.

Agarwal MM, Punnose J, Sukhija K, et al. Gestational diabetes mellitus: using the fasting plasma glucose level to simplify the international association of diabetes and pregnancy study groups diagnostic algorithm in an adult South Asian population. Can J Diabetes 2018;42:500-4.

Mohan V, Multalakshim MM, Bhavdarhuni B, et al. Comparison of screening for gestational diabetes mellitus by oral glucose tolerance tests done in the non-fasting (random) and fasting states. Acta Diabetologica 2014;51:1007-13.

Saxena P, Verma P, Goswami B. Comparison of diagnostic accuracy of non-fasting DIPSI and HbA1c with fasting WHO criteria for diagnosis of gestational diabetes mellitus. J Obstet Gynecol India 2017;67:337-42.

Senanayake H, Seneviratne S, Ariyaratne H, et al. Screening for gestational diabetes mellitus in southern Asian women: screening for gestational diabetes mellitus. J Obstet Gynecol Res 2006;32:286-91.

Sharma M, Nayanisri K, Jain R, et al. Predictive value of fasting plasma glucose on first antenatal visit before 20 weeks of gestation to diagnose gestational diabetes mellitus. J Clin Diagn Res 2018;12:QC01-4.

Siribaddana SH, Deshibandara R, Rajapakse D, et al. The prevalence of gestational diabetes in a Sri Lankan antenatal clinic. Ceylon Med J 1998;43:88-91.

Tripathi R, Verma D, Gupta VK, et al. Evaluation of 75 g glucose load in non-fasting state [Diabetes in Pregnancy Study group of India (DIPSI) criteria] as a diagnostic test for gestational diabetes mellitus. Indian J Med Res 2017;155:209-14.

Wijeyaratne C, Ginge S, Arasalingam A, et al. Screening for gestational diabetes mellitus: the Sri Lankan experience. Ceylon Med J 2009;54:53-8.

Donovan L, Hartling L, Muise M, et al. Screening tests for gestational diabetes: a systematic review for the U.S. preventive services task force. Ann Intern Med 2013;159:115-22.

American College of Obstetricians and Gynecologists Committee on Practice Bulletins-ACOG Practice Bulletin. Clinical management guidelines for obstetrician-gynecologists. Obstet Gynecol 2003;102:647-58.

Huhn EA, Fischer T, Göbl CS, et al. Screening of gestational diabetes mellitus in early pregnancy by oral glucose tolerance test and glycosylated fibronectin: study protocol for an international, prospective, multicentre cohort trial. BMJ Open 2016;6:1-8.

WHO. Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications. Part I: Diagnosis and Classification of Diabetes Mellitus. Geneva: WHO; 1999. 1-59.

Sosniah V, Balaji V, Balaji MS. Scope for prevention of diabetes-focus intrauterine milieu interior. J Assoc Physicians India 2008;56:109-13.

Meltzer S, Snyder J, Penrod J, et al. Gestational diabetes mellitus screening and diagnosis: a prospective randomised controlled trial comparing costs of one-step and two-step methods: costs of gestational diabetes mellitus diagnostic methods: one-step versus two-step. BJOG Int J Obstet Gynecol 2010;117:407-15.

Prutsks GJ, Domecq JP, Sundaresh V, et al. Screening for gestational diabetes: a systematic review and meta-analysis. J Clin Endocrinol Metab 2013;98:4311-8.

Miyakoshi K, Tanaka M, Ueno K, et al. Cutoff value of 1h, 50g glucose challenge test for screening of gestational diabetes mellitus in a Japanese population. Diabetes Res Clin Pract 2003;60:63-7.

Agarwal MM. Gestational diabetes mellitus: screening with fasting plasma glucose. World J Diabetes 2016;7:279-89.

Mishra S, Rao CR, Shetty A. Trends in the diagnosis of gestational diabetes mellitus. Scientific 2016;2016:1-7.

World Health Organization/Diagnostic Criteria and Classification of Hyperglycaemia First Detected in Pregnancy. Geneva: WHO; 2013. 1-62.

The Expert Committee on the Diagnosis and Classification of Diabetes MellitusFollow-up report on the diagnosis of diabetes mellitus. Diabetes Care 2003;26:3160-7.

Carpenter MW, Coustan DR. Criteria for screening tests for gestational diabetes mellitus. J Obstet Gynecol 2010;117:407-15.

National Diabetes Data GroupClassification and diagnosis of diabetes mellitus and other categories of glucose intolerance. Diabetes 1979;28:1039-57.

Triulio J, Vigo A, Reichelt A, et al. Fasting plasma glucose to avoid a full OGTT in the diagnosis of gestational diabetes. Diabetes Res Clin Pract 2014;105:322-6.

World Health OrganizationUse of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus. Diabetes Res Clin Pract 2011;93:299-309.

Khalafallah A, Phuah E, Al-Barazan AM, et al. Glycosylated haemoglobin for screening and diagnosis of gestational diabetes mellitus. BMJ Open 2016;6:1-8.

Renz PB, Cavagnoli G, Wintert LS, et al. Hba1c test as a tool in the diagnosis of gestational diabetes mellitus. PLoS One 2015;10:1-