The accuracy of haemoglobin A1c as a screening and diagnostic test for gestational diabetes: a systematic review and meta-analysis of test accuracy studies

Chiamaka Esther Amaefule (a), Archana Sasitharan (a), Princee Kalra (a), Stamatina Iliodromoti (a), Mohammed S.B. Huda (b), Ewelina Rogozinska (a), Javier Zamora (a,c), and Shakila Thangaratinam (d)

(a) Barts Research Centre for Women's Health, Institute of Population Health Sciences, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK.,
(b) Department of Diabetes & Metabolism, Barts Health NHS Trust, Royal London Hospital, London, UK,
(c) Clinical Biostatistics Unit, Hospital Ramon y Cajal (IRYCIS) and CIBER Epidemiology and Public Health, Madrid, Spain
(d) Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

Correspondence to Chiamaka Esther Amaefule, Barts Research Centre for Women's Health, Institute of Population Health Sciences, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, E1 2AB, UK. Tel: +447931577945;
e-mail: c.e.amaefule@qmul.ac.uk.

Purpose of review

Gestational diabetes mellitus (GDM) is associated with adverse pregnancy complications. Accurate screening and diagnosis of gestational diabetes are critical to treatment, and in a pandemic scenario like coronavirus disease 2019 needing a simple test that minimises prolonged hospital stay. We undertook a meta-analysis on the screening and diagnostic accuracy of the haemoglobin A1c (HbA1c) test in women with and without risk factors for gestational diabetes.

Recent findings

Unlike the oral glucose tolerance test, the HbA1c test is simple, quick and more acceptable. There is a growing body of evidence on the accuracy of HbA1c as a screening and diagnostic test for GDM. We searched Medline, Embase and Cochrane Library and selected relevant studies. Accuracy data for different thresholds within the final 23 included studies (16 921 women) were pooled using a multiple thresholds model. Summary accuracy indices were estimated by selecting an optimal threshold that optimises either sensitivity or specificity according to different scenarios.

Summary

HbA1c is more useful as a specific test at a cut-off of 5.7% (39 mmol/mol) with a false positive rate of 10%, but should be supplemented by a more sensitive test to detect women with GDM.

Keywords

diagnostic test accuracy, gestational diabetes, haemoglobin A1c, meta-analysis, oral glucose tolerance test

KEY POINTS

- HbA1c is commonly used in clinical practice as a screening and diagnostic test for GDM.
- HbA1c is more useful as a specific test (false positive rate of 10%) at a cut-off of 5.7%.
- When sensitivity is optimised, the HbA1c test may have potential but should be interpreted with caution

INTRODUCTION

Hyperglycaemia first diagnosed in pregnancy, known as gestational diabetes, is associated with adverse maternal and foetal outcomes such as preeclampsia, postpartum haemorrhage, still birth and neonatal death [1]. Women with gestational diabetes have up to 70% risk of reoccurrence in subsequent pregnancies and a 50% risk of type-2 diabetes within 5 years after delivery [2,3]. Children born to women with gestational diabetes are at an increased risk of metabolic syndrome denoting a vicious transgenerational cycle of gestational diabetes [4]. Treatment of gestational diabetes improves perinatal outcomes [5]. Early and accurate detection is critical for early intervention with significant public health and economic benefits [6]. There is no consensus on the optimal approach for screening and diagnosing gestational diabetes with variations in clinical practice between and within countries. The two main approaches involve selectively screening pregnant women based on risk factors, an early trimester 50-g glucose challenge test (GCT) or haemoglobin A1c (HbA1c) test in various combinations and subsequently offering a diagnostic 75 or 100 g oral glucose tolerance test (OGTT) at 24–28weeks gestation. This is alternative to a universal approach in which all women are offered an OGTT [7]; however, there is no robust evidence on the long-term benefits on using a universal strategy [6].

The HbA1c test is a measure of glycated haemoglobin which serves as an indicator of glucose control in the prior 2–3 months [8]. Unlike the GCT or OGTT, the HbA1c test is simple, quick and more acceptable to pregnant women as fasting, glucose ingestion and multiple venepunctures are not needed. Prolonged hospital stays due to timed samples are not also required [9]. These characteristics are especially important in the times of infectious diseases pandemic such as coronavirus disease 2019 (COVID-19), were hospital visits and prolonged hospital stays should be kept to absolute minimum to reduce exposure and resultant complications in an already vulnerable group [10&&].

There is a growing body of evidence on the accuracy of first trimester HbA1c as a screening test for gestational diabetes [11–13]. Also, the HbA1c test is commonly used in routine clinical practice as a first trimester screening test for early onset gestational diabetes in women with risk factors for gestational diabetes [14&]. In addition, studies assessing the accuracy of HbA1c test as a diagnostic test for gestational diabetes are increasingly popular [15–17].

We undertook a systematic review and meta-analysis collating evidence on the accuracy of the HbA1c test, both as a first trimester screening test and a 2nd/3rd trimester diagnostic test for gestational diabetes, in women with or without risk factors for gestational diabetes.

METHODS

This review was conducted and reported according to the current guidelines for evidence synthesis of test accuracy [18]. The protocol for this review was registered at PROSPERO (CRD42018080538).

Literature search

Two independent researchers (C.E.A. and P.K.) performed a comprehensive literature search in electronic databases (Medline via Ovid, Embase and Cochrane Library) without any language or time limits. These searches were supplemented with a search for grey literature (OpenGrey). Using search terms specific to HbA1c testing for gestational diabetes, searches were performed on the 23 November 2017 and updated on 18 July 2019. The details of the search strategy can be found in Appendix 1, http://links.lww.com/COOG/A46.

Eligibility criteria and study selection

We included studies with pregnant women with risk factors for gestational diabetes [as per the National Institute for Health and Care Excellence (NICE) guidelines], who underwent the HbA1c test (index test) followed by the OGTT test (reference standard) at the same time point regardless of the

diagnostic criteria applied. As a post-hoc modification to the protocol, we expanded the population to include women without risk factors for gestational diabetes mellitus and studies who administered a first (1st) trimester HbA1c test followed by a second/third (2nd/3rd) trimester OGTT to explore the accuracy of HbA1c in this clinical context. We included prospective and retrospective cohort studies, and excluded case control studies, studies with insufficient data for a 2x2 table, systematic reviews, reports and conference proceedings.

Two independent reviewers (C.E.A. and A.S.) screened the titles and abstracts of citations retrieved from the search, for potentially eligible studies. The full articles of eligible citations were reviewed for their eligibility and included in the review if they fulfilled the selection criteria. Authors of studies with insufficient extractable data to construct a 2x2 table including data on true positive, false positive, true negative and false negative were contacted. Following no response, such studies were excluded from the review and analysis. Consensus on the eligibility of a citation was reached though consultation with a third reviewer (E.R.). The bibliographies of these full text articles were also screened for potentially eligible studies.

Data extraction and study quality assessment

From each selected study, data were extracted onto a pre-piloted data extraction form independently and in duplicate by two reviewers (C.E.A. and A.S.). Information on the following were extracted: study characteristics (first author, year of publication, country, type of cohort study design and the sample size of participants who underwent both the HbA1c test and OGTT); type of population categorised as population with or without risk factors as per the NICE guidelines; reference test characteristics (diagnostic criteria, coded according to its similarity with other criteria in specific cut-offs); index test characteristics (HbA1c thresholds and accuracy parameters) and the trimester of HbA1c and OGTT testing. For studies that have performed the OGTT test using two diagnostic criteria, without stating which criteria has been applied, we have chosen and included the

most recently published criteria in the analysis [17,19]. The data were tabulated and cross-checked consulting a third reviewer (E.R.) in the event of any discrepancies.

All included studies were assessed independently by two reviewers (C.E.A. and A.S.) for risk of bias and applicability using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) checklist tool. We evaluated the four domains including patient selection, implementation of the index and reference tests, the flow and timing of participants. A study was considered to have a low risk of bias if its participants were randomly or consecutively selected, if all participants were tested using the same reference test and if majority of participants were included in the analysis without inappropriate exclusions.

Data synthesis and statistical analysis

We extracted the accuracy parameters (true positive, false positive, true negative and false negative) for all thresholds reported in each individual study. Extracted data were tabulated and summarised by population group (with or without risk factors) and trimester of testing [HbA1c (1st)/OGTT (2nd/3rd) or HbA1c (2nd/3rd)/OGTT 2nd/3rd trimester]. We used the modelling approach proposed by Steinhauser et al. [20] to estimate pooled sensitivity and specificity for an optimal threshold across studies. In short, we used a linear mixed model to estimate the distribution functions of Hb1Ac within the groups of women with and without gestational diabetes assuming a logistic distribution. This approach allows to account for between study heterogeneity and for the correlation between sensitivity and specificity. Model included different random intercepts and different random slopes for women with and without gestational diabetes. We decided the final model specification (intercepts and slopes) guided by REML (Restricted Maximum Livelihood) criterion. From each model we derived the area under the SROC (Summary Receiver Operating Characteristic) curve as a measure of overall accuracy across all thresholds as well as accuracy

indices [sensitivity, specificity, positive likelihood ratio (LR+), negative likelihood ratio (LR-)] for the selected cut-off point.

The primary analysis aimed to optimise sensitivity or specificity based on clinical implication with consideration for population groups (with or without risk factors) and trimester of HbA1c (1st or 2nd/3rd) and OGTT testing. We performed sensitivity analyses assuming normal distribution of Hb1Ac with equal weighting for sensitivity and specificity, also secondarily restricting analysis to studies which have used an OGTT diagnostic criteria of similar cut-offs to the International Association of Diabetes and Pregnancy Study Groups (IADPSG) 2010 criteria. Multiple thresholds analysis was performed using *diagmeta* package in R [21].

RESULTS

A total of 9326 citations were retrieved after a systemic electronic search of the databases. After removing duplicates, 8474 articles were identified and 178 articles selected based on titles and abstract screening. If available, a full text of the article was assessed for eligibility. Twenty-six articles fulfilled the eligibility criteria and their references checked for potentially eligible articles. Two eligible articles were excluded for combining the accuracy estimates from HbA1c testing in the 1st and 2nd/3rd trimester. An additional eligible article was excluded for not stating the trimester of HbA1c or OGTT testing. Twenty-three eligible studies were subsequently included in the review and meta-analysis [11–13,15–17,19,22–37] (Fig. 1).

Characteristics of included studies

The 23 included studies (16,921 pregnant women) were published between 1984 and 2019 and conducted across all economic settings (Table 1). Two population subgroups were included across the 23 studies. Seventeen studies sampled pregnant women with risk factors for gestational diabetes [11–13,15–17,19,22,23,25,26,31,32,34–37] as defined by the NICE guidelines [38] and six

studies included women without risk factors [24,27–30,33]. Studies in settings with an 'at risk' population based on their ethnic origin (South Asians, Africans/Black Caribbean and Middle Eastern) as defined by the NICE guidelines [38] have been classified as having included a population 'at risk' of gestational diabetes even if not explicitly stated or described in the individual studies [16,22,23,35].

Fourteen different diagnostic criteria for the OGTT have been used across all studies [11–13,15,24–26,28–34,36,37]. Six studies have each used two different diagnostic criteria [16,17,19,22,27,35]. Five diagnostic criteria, the IADPSG 2010, modified IADPSG 2010, WHO 2013, American Diabetes Association (ADA) 2013 and The Australasian Diabetes in Pregnancy Society 2013) shared similar cut-offs. Comparably, WHO 1999 and Diabetes in Pregnancy Study group India 2005 as well as modified Carpenter Coustan 1982 and ADA 2004 (Appendix 2, http://links. lww.com/COOG/A47).

The trimester of testing varied across the included studies. In seventeen studies both tests were performed at the same time point; both in the 2nd/3rd trimester [11–13,15-17,22,24,25,27,28, 31,33–37]. In six studies, the HbA1c test was used as a 1st trimester screening test to detect women who may develop gestational diabetes followed by the diagnostic OGTT in the 2nd/3rd trimester [19,26,23,29,30,32]. Included studies reported a range of cut-offs (1–29) with a median of three cut-offs per study (Table 2).

Quality assessment

Based on the four domains of the QUADAS-2 tool, majority of the studies were classified as low risk of bias without major concerns over their applicability. Among studies which included women with risk factors for gestational diabetes, 88% (15/17) were ranked as having a low risk of bias as concerns regarding the interpretation and reporting of the reference standard test criteria led to the classification of high risk of bias in two studies [17,19]. Concerns over applicability were low for all

(17/17) studies who sampled women at risk of gestational diabetes. With studies including women without risk factors, 83% (5/6) were ranked as having a low risk of bias as one study [29] was unclear on the sampling method employed and was assigned an unclear risk of bias in participant selection. There were also no applicability concerns over any study with women without risk factors for gestational diabetes (Fig. 2).

Accuracy of haemoglobin A1c in detecting gestational diabetes in women with risk factors

In the subgroup of women with risk factors for gestational diabetes [17 studies (8067 pregnant women)] and optimising sensitivity at a cut-off of 5.0% (31mmol/mol), the pooled sensitivity and specificity were 0.88 [95% confidence interval (CI) 0.75–0.94] and 0.26 (95% CI 0.15–0.41), respectively. The likelihood ratios of positive and negative test result were 1.18 (95% CI 0.93–1.42) and 0.49 (95% CI 0.05–0.92), respectively (Table 3).

Accuracy of haemoglobin A1c in detecting gestational diabetes in women without risk factors

With the six studies which sampled 8,854 women without risk factors for gestational diabetes, the optimal cut-off was 5.2% (33mmol/mol) with the pooled sensitivity and specificity of 0.86 (95% CI 0.47–0.98) and 0.32 (95% CI 0.06–077), respectively. The LR+ and LR- were 1.28 (95% CI 0.40–2.15) and 0.43 (95% CI 0.00–1.34), respectively (Table 3).

Accuracy of haemoglobin A1c as a 1st trimester screening test for gestational diabetes

Six studies (7100 women) administered the HbA1c test in the 1st trimester as a screening test for gestational diabetes and subsequently the OGTT in the 2nd/3rd trimester. With the aim of maximizing sensitivity, the pooled sensitivity and specificity were 0.93 (95% CI 0.66–0.99) and 0.22 (95% CI 0.05–0.62), respectively. The optimal cut-off was 5.2% (33 mol/mol). The LR+ and LR- for a first trimester HbA1c test to predict the onset of gestational diabetes by the 2nd/3rd trimester were 1.18 (95% CI 0.71–1.66) and 0.34 (95% CI 0.00–1.08), respectively (Table 3).

Accuracy of haemoglobin A1c as a 2nd/3rd trimester diagnostic test for gestational diabetes

With the seventeen studies (9821 women) which assessed the diagnostic accuracy of the HbA1c test as a 2nd/3rd trimester diagnostic test for gestational diabetes, the pooled sensitivity was 0.82 (95% CI 0.70–0.89) and pooled specificity 0.40 (95% CI 0.29–0.54) at a slightly lower cut-off of 5.1% (32mmol/mol). The LR+ is 1.37 (95% CI 1.04–1.71) and LR- 0.45 (95% CI 0.18–0.73) (Table 3).

If specificity is optimised, the pooled sensitivity and specificity of the HbA1c test as a 2nd/3rd trimester diagnostic test was 0.36 (95% CI 0.23–0.52) and 0.90 (95% CI 0.79–0.95) at a cut-off of 5.7% (39mmol/mol). The LR+ was 3.55 (95% CI 0.51–6.58) and the LR- was 0.71 (95% CI 0.53–0.89). In the subgroup of women with risk factors (n¼6767), the cut-off is slightly higher at 5.9% (41mmol/mol) with slight changes in pooled sensitivity 0.35 (95% CI 0.20–0.53) and pooled specificity 0.91 (95% CI 0.78–0.97). The LR+ was slightly higher at 3.77 (95% CI 0.00–7.75) whereas LR- was similar 0.72 (95% CI 0.52–0.92).

Sensitivity analysis

If we aim to optimise both sensitivity and specificity, including all twenty-three studies, we found poor sensitivity of 0.56 (95% CI 0.39–0.72) and a specificity of 0.72 (0.49–0.88) at a cut-off of 5.5% (37mmol/mol). A sensitivity analysis, pooling studies with the same cut-off as the IADPSG 2010 diagnostic criteria yielded a pooled sensitivity of 0.58 (95% CI 0.44–0.71) and pooled specificity of 0.71 (95% CI 0.60–0.80), at an optimal cut-off of 5.3% (34mmol/mol).

DISCUSSION

To avoid a missed diagnosis and resultant adverse clinical complications of gestational diabetes [1], a highly sensitive test with low false negative rate will be ideal to enable the identification of cases and

immediate commencement of appropriate treatment and management approaches. This is particularly important in low-income settings where access to antenatal services and additional safety nets such as frequent urine tests, serial growth scans or further antenatal checks which could detect the onset of gestational diabetes at a later gestation, is limited [39]. The clinical implications of a missed diagnosis informed our decision to optimise the sensitivity of the HbA1c test to determine the optimal cut-off and accuracy parameters.

As expected, the HbA1c test has a high sensitivity when sensitivity is optimised, but a poor overall performance in detecting gestational diabetes in pregnant women regardless of their risk and trimester of testing. The likelihood of a positive test result to rule in gestational diabetes is also poor. However, in women with risk factors, the false negative rate is lower (12%) at a 5.0% cut-off than in women without risk factors. The false negative rate is also lower when HbA1c is used as a first trimester screening test (false negative rate of 7% at a 5.2% cut-off) rather than a 2nd/3rd trimester diagnostic test, but still at a fairly low false negative rate when used as a diagnostic test (false negative rate of 18% at 5.1% cut-off).

Although with a low false negative rate, the trade-off of a high false positive rate ranging from 60 to 78% results in unnecessary treatments which can further drain a low resource health system as a high proportion of women with a positive test will not develop gestational diabetes. In addition, over burden an already strained health workforce like in a pandemic era [10&&]. However, the consequence of missing a diagnosis has more severe clinical implications than over treatment as mild to moderate levels of hyperglycaemia are managed with dietary measures which in turn have other health benefits [38].

On the other hand, the HbA1c test can be useful as a specific test in clinical scenarios requiring a first trimester screening test or a triage test to minimise false positives: unnecessary OGTTs in women

who will not develop gestational diabetes given the challenge with acceptability and to allow the efficient use of limited resources towards a confirmatory 2nd/3rd trimester test in low resource health systems. In addition, HbA1c can be useful as a highly specific diagnostic test in an infectious diseases pandemic era like COVID-19 to minimise unnecessary exposure in an already vulnerable group preventing resultant complications, but also avoid burdening an already drained workforce. In addition, the nature of the HbA1c test as a simple, quick test not requiring timed samples or prolonged hospital stays minimises the risk of transmission to frontline healthcare staff and also saves valuable time.

When specificity is optimised, the HbA1c test performs better at ruling in gestational diabetes with a higher LR+ when used as a 2nd/3rd trimester diagnostic test in women with or without risk factors for gestational diabetes. At a high specificity, the false positive rate is minimised at 10% with a cut-off of 5.7% (39mmol/mol) and further at 9% at a cut-off of 5.9% (41mmol/mol cut-off) in women with risk factors. However, with high specificity is a concomitant reduction in sensitivity resulting in a high false negative rate, 0.64 and 0.65 respectively. To compensate for this poor sensitivity, this high specific HbA1c cut-off of 5.7% (39mmol/mol) should be used in combination with a highly sensitive, simple and acceptable test to increase the odds of detecting women with gestational diabetes. These findings are consistent with a previous review [40] and the ongoing MRC (Medical Research Council) funded PRegnancy and Infant Development study [41] on the usefulness of the HbA1c 5.7% (39mmol/mol) cut-off to diagnose gestational diabetes to minimise the false positive rate especially in a pandemic era, but supplemented by a more sensitive test. This in turn informed the Royal College of Obstetricians and Gynaecologists guideline on the HbA1c cut-off of 39mmol/mol as the alternative threshold for diagnosing gestational diabetes in the COVID-19 pandemic era [10&&]. There are strengths to this review. Unlike a previous review [40], this systematic review reports on the accuracy of HbA1c both as a first trimester screening test and a 2nd/3rd trimester diagnostic test, also reporting on population-specific cut-offs. Methodologically, instead of eight studies, we

have included 23 studies (16,921 women) of predominantly high quality without any time or language restrictions. As an additional strength, rather than reporting on accuracy measures for each threshold, our statistical approach takes into account all the cut-offs reported by individual studies in deriving the pooled sensitivity and specificity of the screening and diagnostic accuracy of the HbA1c test.

Although there was heterogeneity between studies due to the variation in population groups sampled, trimester of HbA1c testing and the OGTT diagnostic criteria employed, we used a linear mixed model which accounts for between study heterogeneity and correlation between sensitivity and specificity. As a limitation, about one-third of the eligible citations based on title and abstract screening where not available by full text and could not be assessed for inclusion, even though librarian services and request to authors. In addition, our decision to pool studies with different trimester of testing or risk groups can be a limitation, but the number of studies in certain groups (such as the numbers of studies with population group without risk factors) precludes us from making further stratifications. Secondly, although test performance is not affected by prevalence, in our sub-group analyses on the accuracy of HbA1c as a 2nd/3rd diagnostic test regardless of population groups, the optimal cut-offs and accuracy parameters vary only slightly to the diagnostic performance of the test in only women with risk factors (Table 3).

CONCLUSION

Our systematic review showed that HbA1c is useful as a specific test to rule-out gestational diabetes. In certain clinical scenarios requiring to minimize false positives, we provide optimal cut-offs which can be useful but at the expense of a missed diagnosis which in turn can be detected when supplemented by a more sensitive test. If sensitivity is optimised, HbA1c may have potential in identifying cases using the optimal cut-offs reported in this study, but should be interpreted with caution.

Acknowledgements

Author contributions: C.E.A. wrote the protocol, conducted the literature search, study selection, data extraction, quality assessment, interpreted and reported on the statistical findings. C.E.A. wrote the first and final draft of the article. A.S. and P.K. duplicated the literature search, study selection and data extraction as a second reviewer and have approved the final article. S.I. and M.S.B.H. provided clinical and methodological input and contributed to final draft of the article. E.R. supervised the review process and contributed to the first and final draft of the article. J.Z. conducted the statistical analysis and contributed to the final draft of the article. S.T. conceived the research question, supervised the review process and contributed to the final draft of the article.

Financial support and sponsorship

None.

Conflicts of interest

There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

& of special interest

&& of outstanding interest

1. Metzger BE, Lowe LP, Dyer AR, et al. Hyperglycemia and adverse pregnancy outcomes. N Engl J Med 2008; 358:1991–2002.

2. Major CA, deVeciana M, Weeks J, Morgan MA. Recurrence of gestational diabetes: who is at risk? Am J Obstet Gynecol 1998; 179:1038–1042.

3. Bellamy L, Casas JP, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. Lancet 2009; 373:1773–1779.

4. Boney CM, Verma A, Tucker R, Vohr BR. Metabolic syndrome in childhood: association with birth weight, maternal obesity, and gestational diabetes mellitus. Pediatrics 2005; 115:e290–e296.

5. Crowther CA, Hiller JE,Moss JR, et al. Effect of treatment of gestational diabetes mellitus on pregnancy outcomes. N Engl J Med 2005; 352:2477–2486.

6. Farrar D, Simmonds M, Griffin S, et al. The identification and treatment of women with hyperglycaemia in pregnancy: an analysis of individual participant data, systematic reviews, meta-analyses and an economic evaluation. Health Technol Assess 2016; 20:1–348.

7. Metzger BE, Gabbe SG, Persson B, et al., International Association of Diabetes and Pregnancy Study Groups Consensus Panel. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care 2010; 33:676–682.

8. Nathan DM, Turgeon H, Regan S. Relationship between glycated haemoglobin levels and mean glucose levels over time. Diabetologia 2007; 50:2239–2244.

9. Agarwal MM, Punnose J, Dhatt GS. Gestational diabetes: problems associated with the oral glucose tolerance test. Diabetes Res Clin Pract 2004; 63:73–74.

10. && Royal College of Obstetricians and Gynaecologists. Guidance for maternal medicine in the evolving coronavirus (COVID-19) pandemic. In: Information for healthcare professionals. ENGLAND: RCOG and NHS; 2020:; Web site: https://www.rcog.org.uk/globalassets/documents/guidelines/2020-03-30-guidance-for-maternal-medicine-in-the-evolving-coronavirus-covid-19-pandemic.pdf. [Published 2020; Accessed 4 April] Royal College of Obstetricians and Gynaecologists guideline on the usefulness of HbA1c as an alternative test in the diagnosis of gestational diabetes in the COVID- 19 pandemic era.

11. Agarwal MM, Dhatt GS, Punnose J, Koster G. Gestational diabetes: a reappraisal of HBA1c as a screening test. Acta Obstet Gynecol Scand 2005; 84:1159–1163.

12. Khalafallah A, Phuah E, Al-Barazan AM, et al. Glycosylated haemoglobin for screening and diagnosis of gestational diabetes mellitus. BMJ Open 2016; 6:e011059.

13. Odsaeter IH, Asberg A, Vanky E, Carlsen SM. HbA1c as screening for gestational diabetes mellitus in women with polycystic ovary syndrome. BMC Endocr Disord 2015; 15:38.

14. & Sanghi AH, Saravanamuthu SBM, Jamna T, et al. Maternity services in diabetes–pregnancy, labour and puerperium. In: Barts NHS Trust Local Policy guidelines. Barts Health NHS Trust: 2019. Barts NHS Trust Policy on the use of HbA1c as a first trimester screening test for early onset gestational diabetes in women with risk factors for this condition.

15. Agarwal MM, Hughes PF, Punnose J, et al. Gestational diabetes screening of a multiethnic, high-risk population using glycated proteins. Diabetes Res Clin Pract 2001; 51:67–73.

16. Rajput R, Yogesh Yadav. Rajput M, Nanda S. Utility of HbA1c for diagnosis of gestational diabetes mellitus. Diabetes Res Clin Pract 2012; 98: 104–107.

17. Renz PB, Cavagnolli G, Weinert LS, et al. HbA1c test as a tool in the diagnosis of gestational diabetes mellitus. PLoS One 2015; 10:e0135989.

18. McInnes MDF, Moher D, Thombs BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. JAMA 2018; 319:388–396.

19. 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; 211:641.e1–641.e7.

20. Steinhauser S, Schumacher M, R€ucker G. Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies. BMC Med Res Methodol 2016; 16:97.

21. Rucker GSS, Kolampally S, Schwarzer G. Diagmenta: meta-analysis of diagnostic accuracy studies with several cut-off points, R package version 0.3-0. Guido Schwarzer; 2018; https://CRAN.R-project.org/package=diagmeta [Accessed 21 April 2020]

406 22. Saxena P, Verma P, Goswami B. comparison of diagnostic accuracy of nonfasting DIPSI and
407 HbA1c with fasting WHO criteria for diagnosis of gestational diabetes mellitus. J Obstet Gynaecol
408 India 2017; 67:337–342.
409 23. Arbib N, Shmueli A, Salman L, et al. First trimester glycosylated hemoglobin as a predictor of
410 gestational diabetes mellitus. Int J Gynaecol Obstet 2019; 145:158–163.
411 24. Siricharoenthai P, Phupong V. Diagnostic accuracy of HbA1c in detecting gestational diabetes
412 mellitus. J Matern Fetal Neonatal Med 2019; 11:1–4.
413 25. Soumya S, Rohilla M, Chopra S, et al. HbA1c: a useful screening test for gestational diabetes
414 mellitus. Diabetes Technol Ther 2015; 17:899–904.
415 26. Amylidi S, Mosimann B, Stettler C, et al. First-trimester glycosylated haemoglobin in women at
416 high risk for gestational diabetes. Acta Obstet Gynecol Scand 2016; 95:93–97.
417 27. Odsaeter IH, Asberg A, Vanky E, et al. Hemoglobin A1c as screening for gestational diabetes
418 mellitus in Nordic Caucasian women. Diabetol Metab Syndr 2016; 8:43.
419 28. Ye M, Liu Y, Cao X, et al. The utility of HbA1c for screening gestational diabetes mellitus and its
420 relationship with adverse pregnancy outcomes. Diabetes Res Clin Pract 2016; 114:43–49.
421 29. Benaiges D, Flores-Le Roux JA, Marcelo I, et al. Is first-trimester HbA1c useful in the diagnosis of
422 gestational diabetes? Diabetes Res Clin Pract 2017; 133:85–91.
423 30. Hughes RCE, Moore MP, Gullam JE, et al. An early pregnancy HbA1c _5.9% (41 mmol/mol) is
424 optimal for detecting diabetes and identifies women at increased risk of adverse pregnancy
425 outcomes. Diabetes Care 2014; 37:2953–2959.
426 31. Sevket O, Sevket A, Ozel A, et al. The use of HbA1c as an aid in the diagnosis of gestational
427 diabetes mellitus. J Obstet Gynaecol 2014; 34:690–692.
428 32. Osmundson SS, Zhao BS, Kunz L, et al. First trimester hemoglobin A1c prediction of gestational
429 diabetes. Am J Perinatol 2016; 33:977–982.

33. Ryu AJ, Moon HJ, Na JO, et al. The usefulness of the glycosylated haemoglobin level for the diagnosis of gestational diabetes mellitus in the Korean population. Diabetes Metab J 2015; 39:507–511.

34. Ho YR, Wang P, Lu MC, et al. Associations of mid-pregnancy HbA1c with gestational diabetes and risk of adverse pregnancy outcomes in high-risk Taiwanese women. PLoS One 2017; 12:e0177563.

35. Agbozo F, Abubakari A, Narh C, Jahn A. Are we missing pregnant women with gestational diabetes? Evidence from a diagnostic accuracy study comparing glycosuria, glycated haemoglobin, random and fasting glucose to oral glucose tolerance test. Trop Med Int Health 2017; 22:351–352.

36. Veres¸ M, Cr3iut¸ DI, Trutz J, Babes¸ A. The utility of glycated hemoglobin, determined in the second trimester of pregnancy, in diagnosing gestational diabetes. Rom J Diab Nutr Metab Dis 2015; 22:233–240.

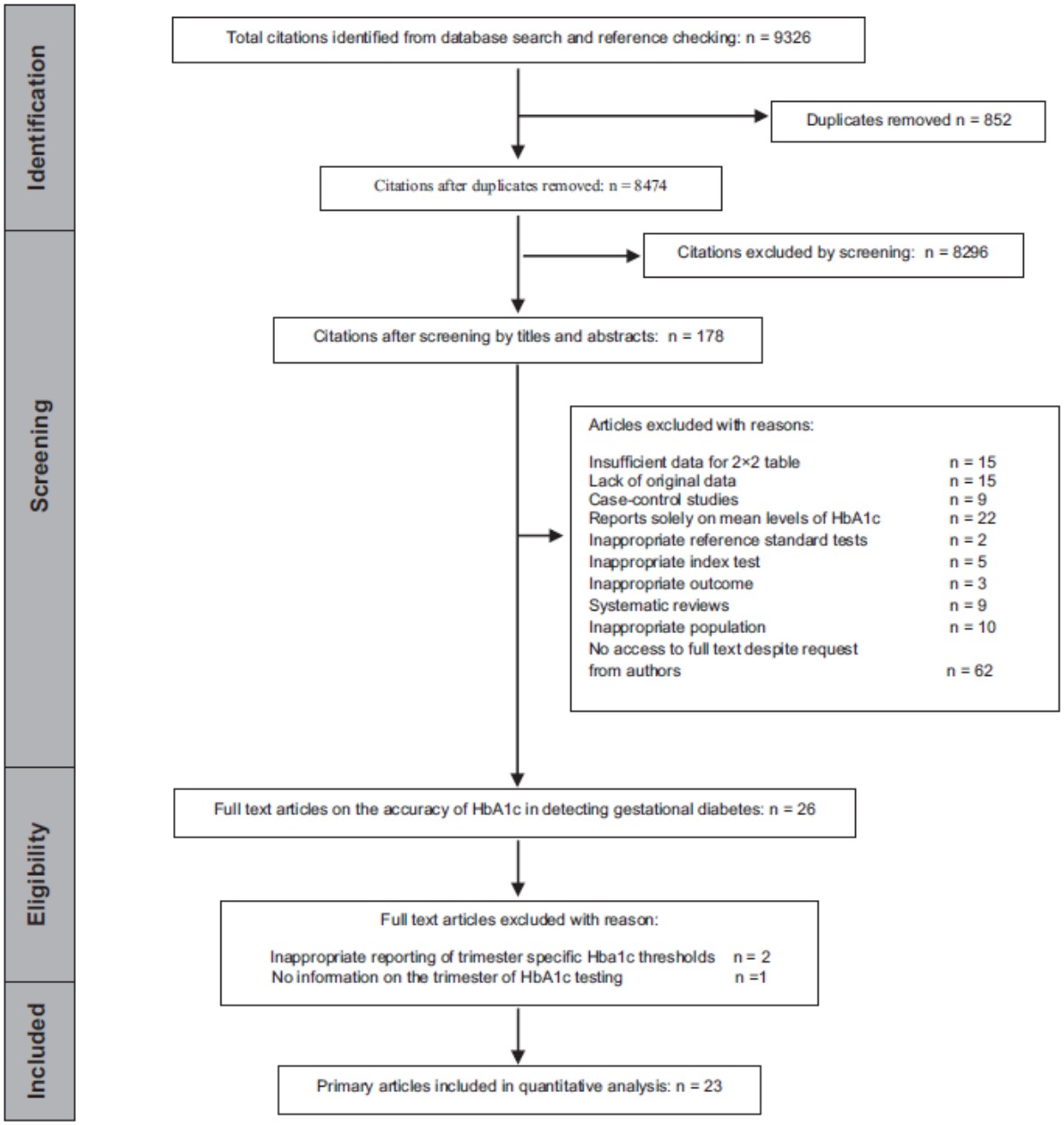
37. Artal R, Mosley GM, Dorey FJ. Glycohemoglobin as a screening test for gestational diabetes. Am J Obstet Gynecol 1984; 148:412–414.

38. National Institute for Health and Care Excellence. Diabetes in pregnancy: management from preconception to the postnatal period. In: NICE guideline [NG3]. NICE; 2019; Web site: https://www.nice.org.uk/guidance/ng3/chapter/1-Recommendations#gestational-diabetes-2. [Published 2015; Accessed 26 November 2019]

39. Campbell OM, Graham WJ; Lancet Maternal Survival Series steering group. Strategies for reducing maternal mortality: getting on with what works. Lancet 2006; 368:1284–1299.

40. Renz PB, Chume FC, Timm JRT, et al. Diagnostic accuracy of glycated hemoglobin for gestational diabetes mellitus: a systematic review and meta-analysis. Clin Chem Lab Med 2019; 57:1435–1449.

41. van Gelder MM, Bretveld RW, Roukema J, et al. Rationale and design of the PRegnancy and Infant DEvelopment (PRIDE) Study. Paediatr Perinat Epidemiol 2013; 27:34–43.



454

455 FIGURE 1. Flow diagram showing the selection of diagnostic test accuracy studies comparing

456 haemoglobin A1c and oral glucose tolerance test in detecting gestational diabetes.

457

458

459 Table 1. Characteristics of studies included in the systematic review of the accuracy of haemoglobin

460 A1c in detecting gestational diabetes

Study ID	Country	Design (type of cohort study)	Sample size	Population	Diagnostic criteria	Trimester of OGTT testing	Trimester of HbA1c testing
Agarwal, 2001	United Arab Emirates,	Prospective	426	with risk factors	Modified CC 1982	2nd/3rd	2nd/3rd
Agarwal, 2005	United Arab Emirates	Prospective	442	with risk factors	WHO 1999	2nd/3rd	2nd/3rd
Agbozo, 2017	Ghana	Prospective	480	with risk factors	WHO 2013 & NICE 2015	2nd/3rd	2nd/3rd
Amylidi, 2015	Switzerland	Retrospective	218	with risk factors	ADA 2015	1st	2nd/3rd
Artal, 1984	California	Prospective	82	with risk factors	ADA 1980	3rd	3rd
Benaiges, 2017	Barcelona	Retrospective	1158	without risk factors	NDDG 1979	1st	2nd/3rd
Fong, 2014	California	Retrospective	526	with risk factors	CC 1982 & IADPSG 2010	1st/2nd	2nd/3rd
Ho, 2017	Taiwan	Prospective	1989	with risk factors	CC 1982	2nd/3rd	2nd/3rd
Hughes, 2014	New Zealand	Prospective	4642	without factors	NZ 2014	1st/2nd	2nd/3rd
Khalafallah, 2016	Australia	Prospective	480	with risk factors	ADIPS 2013	2nd/3rd	2nd/3rd
Odsaeter, 2015	Norway	Prospective	228	with risk factors	WHO 1999	$1^{st}/2^{nd}$/3rd	$1^{st}/2^{nd}$/3rd

Odsaeter, 2016	Norway	Retrospective	638	without risk factors	WHO 1999 & Modified IADPSG 2010	2nd/3rd	2nd/3rd
Osmundon, 2016	Northern California, USA	Retrospective	414	with risk factors	IADPSG 2010	1st/2nd	2nd/3rd
Rajput, 2012	India	Prospective	607	without risk factors	ADA 2004 & IADPSG 2010	2nd/3rd	2nd/3rd
Renz, 2015	Brazil	Prospective	262	with risk factors	WHO 2013 & WHO 1999	2nd/3rd	2nd/3rd
Ryu, 2015	South Korea	Retrospective	343	without risk factors	CC 1982	2nd/3rd	2nd/3rd
Saxena, 2017	India	Prospective	800	With risk factors	DIPSI 2005 & WHO 199	2nd/3rd	2nd/3rd
Sevket, 2014	Turkey	Prospective	339	with risk factors	IADPSG 2010	2nd/3rd	2nd/3rd
Siricharoenthai, 2019	Thailand	Prospective	114	Without risk factors	NDDG 1979	2nd/3rd	2nd/3rd
Soumya, 2015	India	Prospective	500	without risk factors	IADPSG 2010	2nd/3rd	2nd/3rd
Veres, 2015	Romania	Prospective	132	with risk factors	CC 1982	2nd/3rd	2nd/3rd
Ye, 2016	China	Retrospective	1959	without risk factors	IADPSG 2010	2nd/3rd	2nd/3rd

461 ADA, American Diabetes Association; ADIPS, The Australasian Diabetes in Pregnancy Society; BDA,
462 British Diabetic Association; CC, Carpenter Coustan; CMA, Chinese Medical Association; DIPSI,
463 Diabetes in Pregnancy Study Group India; HbAlc, haemoglobin A1c; IADPSG, International
464 Association of Diabetes and Pregnancy Study Groups; NDDG, National Diabetes Data Group; NICE,
465 The National Institute for Health and Care Excellence; NZ, New Zealand Guidelines; OGTT, oral
466 glucose tolerance test.

467

468 Table 2 Accuracy of the haemoglobin A1c test for all thresholds extracted from individual studies

Study ID	Sample size (HbA1c and OGTT)	Trimester of HbA1c testing	Trimester of OGTT testing	Diagnostic criteria	Hba1c threshold ≥ (%)	TP	TN	FP	FN
Agarwal, 2001	426	2nd/3rd	2nd/3rd	Modified CC. 1982	4.5	112	14	298	2
					5	105	86	226	9
					5.5	83	206	106	31
					6	39	284	28	75
					6.5	15	306	6	99
					7	4	311	1	110
Agarwal, 2005	442	2nd/3rd	2nd/3rd	WHO 1999	4.5	82	5	353	2
					5	82	17	341	2
					5.5	69	75	283	15
					6	41	199	159	43
					6.5	18	281	77	66
					7	9	324	34	75
					7.5	6	343	15	78
					8	3	353	5	81
Agbozo, 2017	480	2nd/3rd	2nd/3rd	WHO 2013	6.5	1	364	26	37
				NICE 2015	6.5	2	342	25	59
Amylidi, 2016	218	1st	2nd/3rd	ADA 2015	5.25	24	95	91	8
Arbib, 2018	142	1st	2nd/3rd	CC 1982	5.45	35	69	31	7
Artal, 1984	82	3rd	3rd	ADA 1980	7	22	17	33	8
Benaiges, 2017	1158	1st	2nd/3rd	NDDG 1979	4.5	151	24	982	1
					4.6	150	42	964	2
					4.7	149	67	939	3
					4.8	147	102	904	5
					4.9	141	180	826	11
					5	129	273	733	23
					5.1	120	399	607	32
					5.2	111	540	466	41
					5.3	98	646	360	54
					5.4	82	750	256	70
					5.5	67	834	172	85
					5.6	50	898	108	102
					5.7	39	931	75	113
					5.8	30	955	51	122
					5.9	22	981	25	130
					6	16	992	14	136
					6.1	11	1000	6	141
Fong, 2014	526	1st/2nd	2nd/3rd	CC 1982 & IADPSG 2010	5.7	15	430	40	41
Ho, 2017	1989	2nd/3rd	2nd/3rd	CC 1982	5.7	260	1188	225	316
Hughes, 2014	4642	1st/2nd	2nd/3rd	NZ 2014	5.9	568	71	3863	140
					6.5	16	3942	2	692
Khalafallah, 2016	480	2nd/3rd	2nd/3rd	ADIPS 2013	4.6	55	19	404	2
					4.7	55	42	381	2
					4.8	47	76	347	10
					4.9	42	133	290	15

					5	40	220	203	17
					5.1	35	286	137	22
					5.2	31	337	86	26
					5.3	20	374	49	37
					5.4	15	404	19	42
					5.5	13	415	8	44
					5.6	7	419	4	50
					5.7	6	421	2	51
					5.8	5	422	1	52
					5.9	3	422	1	54
					6	2	422	1	55
					6.1	1	422	1	56
					10	0	422	1	57
Odsaeter, 2015	228	1st	1st	WHO 1999	4.7	20	1	207	0
					4.8	19	5	203	1
					4.9	19	16	192	1
					5.4	6	195	13	14
					5.5	5	203	5	15
					5.6	4	208	0	16
		2nd/3rd	2nd/3rd	WHO 1999	4.7	55	1	172	0
					4.8	54	5	168	1
					4.9	53	15	158	2
					5.4	8	162	11	47
					5.5	5	168	5	50
					5.6	4	173	0	51
Odsaeter, 2016	638			WHO 1999	4.7	37	22	579	0
					4.8	36	48	553	1
					4.9	34	103	498	3
					5.5	11	571	30	26
					5.6	8	587	14	29
					5.9	0	601	0	37
	629	2nd/3rd	2nd/3rd	WHO 1999	4.4	42	3	584	0
					4.5	41	14	573	1
					4.6	40	45	542	2
					5.2	6	559	28	36
					5.3	4	574	13	38
					5.8	0	587	0	42
	677	2nd	2nd	Modified IADPSG 2010	4.7	16	110	551	0
					4.8	14	200	461	2
					4.9	11	339	322	5
					5.2	2	626	35	14
					5.3	1	644	17	15
	627	3rd	3rd	Modified IADPSG 2010	4.6	29	6	592	0
					5	28	188	410	1
					5.1	24	288	310	5
					5.5	9	566	32	20
					5.6	6	583	15	23
					5.8	2	598	0	27
	628	2nd/3rd	2nd/3rd	Modified IADPSG 2010	4.4	45	3	580	0
					4.6	44	14	569	1
					4.7	43	96	487	2
					5.2	6	555	28	39
					5.3	4	570	13	41
					5.8	0	583	0	45

					5.7	54	2254	144	360
Osmundson, 2016	414	$1^{st}/2^{nd}$	$1^{st}/2^{nd}$	IADPSG 2010	5.7	54	2254	144	360
Rajput, 2012	607	$2^{nd}/3^{rd}$	$2^{nd}/3^{rd}$	ADA 2004	5.45	37	345	219	6
					5.95	12	548	16	31
				IADPSG 2010	5.25	120	188	275	24
					5.95	17	450	13	127
Renz, 2015	262	$2^{nd}/3^{rd}$	$2^{nd}/3^{rd}$	WHO 2013 & WHO 1999	5	77	57	119	9
					5.1	72	78	98	14
					5.2	67	103	73	19
					5.3	60	118	58	26
					5.4	54	134	42	32
					5.5	44	146	30	42
					5.6	36	155	21	50
					5.7	27	160	16	59
					5.8	23	167	9	63
					5.9	18	171	5	68
					6	13	173	3	73
					6.5	6	176	0	80
Ryu, 2015	343	$2^{nd}/3^{rd}$	$2^{nd}/3^{rd}$	CC 1982	5.05	107	37	197	2
					5.15	104	70	164	5
					5.25	102	102	132	7
					5.35	95	166	68	14
					5.45	80	198	36	29
					5.55	55	211	23	54
					5.65	46	224	10	63
					5.75	34	228	6	75
					5.85	26	232	2	83
Saxena, 2017	800	$2^{nd}/3^{rd}$	$2^{nd}/3^{rd}$	WHO 1999	6	24	733	16	27
Sevket, 2014	339	$2^{nd}/3^{rd}$	$2^{nd}/3^{rd}$	IADPSG 2010	4.6	51	66	220	2
					5.2	34	193	93	19
					5.7	14	259	27	39
Siricharoenthai, 2019	114	$2^{nd}/3^{rd}$	$2^{nd}/3^{rd}$	NDDG 1979	5.8	6	79	0	29
Soumya, 2015	500	$2^{nd}/3^{rd}$	$2^{nd}/3^{rd}$	IADPSG 2010	5.3	43	232	223	2
					5.7	33	344	111	12
					6.1	21	432	23	23
Veres, 2015	132	$2^{nd}/3^{rd}$	$2^{nd}/3^{rd}$	CC 1982	5.1	26	84	22	0
					5.7	15	97	9	11
					5.8	5	88	18	21
					6.5	12	103	3	14
					6.8	11	104	2	15
					7	5	99	7	21
Ye, 2016	1959	$2^{nd}/3^{rd}$	$2^{nd}/3^{rd}$	IADPSG 2010	4.1	413	22	1524	0
					4.2	412	32	1514	1
					4.3	410	53	1493	3
					4.4	409	83	1463	4
					4.5	400	141	1405	13
					4.6	388	220	1326	25
					4.7	366	337	1209	47
					4.8	351	492	1054	62
					4.9	307	683	863	106
					5	267	900	646	146

					5.1	204	1090	456	209
					5.2	172	1260	286	241
					5.3	126	1354	192	287
					5.4	80	1436	110	333
					5.5	61	1480	66	352
					5.6	46	1510	36	367
					5.7	37	1531	15	376
					5.8	26	1538	8	387
					5.9	20	1544	2	393
					6	13	1544	2	400
					6.1	10	1544	2	403
					6.2	7	1546	0	406

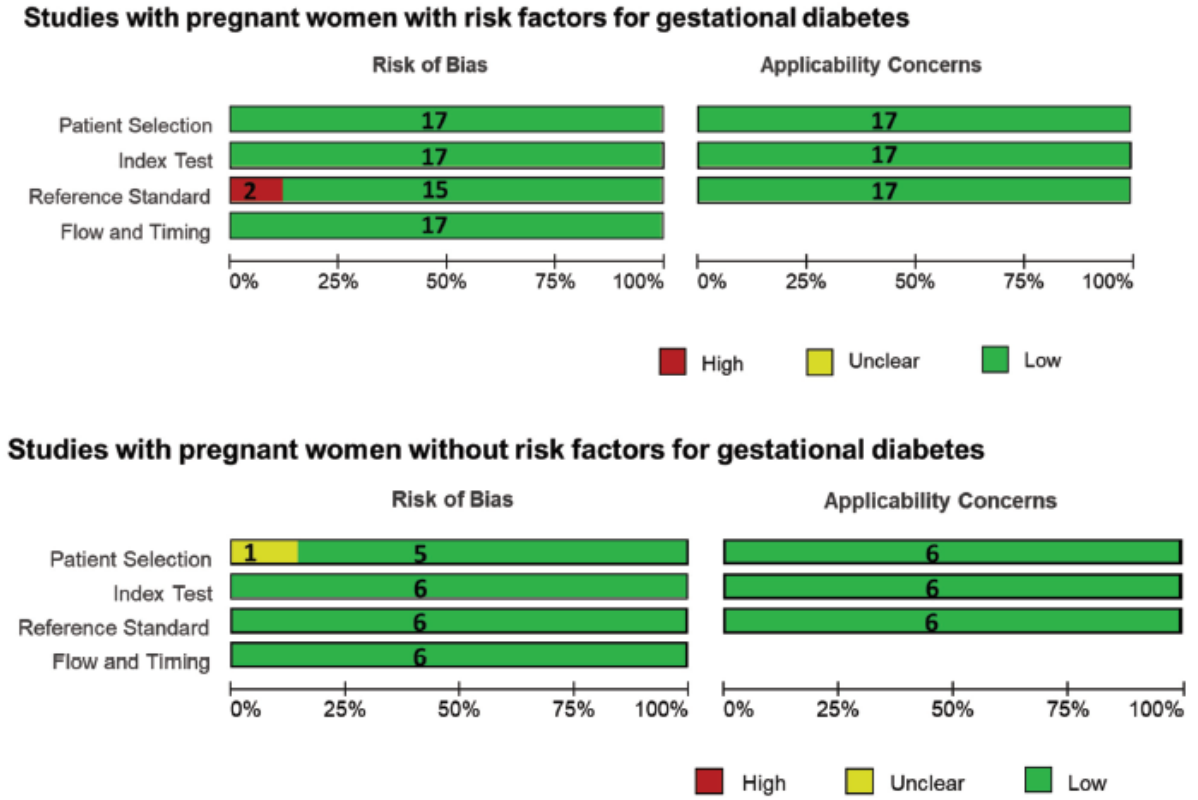


FIGURE 2 Quality assessment of included studies (using Quality Assessment of Diagnostic Accuracy Studies-2 tool) on the accuracy of haemoglobin A1c testing to detect gestational diabetes in women with or without risk factors for gestational diabetes.

Table 3 Accuracy of the haemoglobin A1c test as a screening or diagnostic test for gestational diabetes in women with or without risk factors

Optimising sensitivity [a]				
	Women with risk factors	Women without risk factors	All women	
	1st and 2nd / 3rd trimester		1st trimester	2nd / 3rd trimester
Sensitivity (95% CI)	0.88 (0.75-0.94)	0.86 (0.47-0.98)	0.93 (0.66-0.99)	0.82 (0.70-0.89)
Specificity (95% CI)	0.26 (0.15-0.41)	0.32 (0.06-0.77)	0.22 (0.05-0.62)	0.40 (0.29-0.54)
LR+ (95% CI)	1.18 (0.93-1.42)	1.28 (0.40-2.15)	1.18 (0.71-1.66)	1.37 (1.04-1.71)
LR- (95% CI)	0.49 (0.05-0.92)	0.43 (0.00-1.34)	0.34 (0.00-1.08)	0.45 (0.18-0.73)
Optimal cut-off (%)	5.0	5.2	5.2	5.1
HbA1c value (mmol/mol)	31	33	33	32
AUC	0.67	0.69	0.63	0.71
Number of studies	17	6	6	17

Optimising specificity			Equal weighting	
	Women with risk factors	All women	All women	
	2nd / 3rd trimester		1st and 2nd / 3rd trimester	
Sensitivity (95% CI)	0.35 (0.20-0.53)	0.36 (0.23-0.52)	0.56 (0.39-0.72)	0.58 (0.44-0.71)
Specificity (95% CI)	0.91 (0.78-0.97)	0.90 (0.79-0.95)	0.72 (0.49-0.88)	0.71 (0.60-0.80)
LR+ (95% CI)	3.77 (0.00-7.75)	3.55 (0.51-6.58)	2.03 (0.42-3.64)	2.03 (1.18-2.89)
LR- (95% CI)	0.72 (0.52-0.92)	0.71 (0.53-0.89)	0.61 (0.32-0.89)	0.59 (0.37-0.80)
Optimal cut-off (%)	5.9	5.7	5.5	5.3
HbA1c value (mmol/mol)	41	39	37	34
AUC	0.70	0.71	0.68	0.69
Number of studies	13	17	23	11 [b]

AUC, area under the curve; CI, confidence interval; FN, false negative; FP, false positive; IADPSG, International Association of Diabetes and Pregnancy Study Groups; HbA1c, haemoglobin A1c; LR+, positive likelihood ratio; LR-, negative likelihood ratio.

[a] Costs for FN=2xFP – the health economics cost of a false negative test is twice the cost of a false positive test.

[b] Studies with similar cut-offs to the IADPSG 2010 diagnostic criteria