Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Performance of guidelines for the screening and diagnosis of gestational diabetes mellitus during the COVID-19 pandemic: A scoping review of the guidelines and diagnostic studies evaluating the recommended testing strategies

Aisling M. Curtis, Andrew J. Farmer, Nia W. Roberts, Laura C. Armitage

Purpose: The COVID-19 pandemic has necessitated less resource-intensive testing guidelines to identify gestational diabetes mellitus (GDM). We performed a scoping review of the international evidence reporting the ability of diagnostic tests recommended during the pandemic to accurately identify patients with GDM, compared to pre-pandemic reference standards, and associated test and clinical outcomes.

Methods: A comprehensive search of the literature was carried out in Embase, LitCovid, Cochrane Covid-19 study register, and medRxiv on 14th June 2021.

Results: 145 unique citations were returned; after screening according to pre-specified inclusion criteria by title and abstract and then full text, 13 studies involving 40,836 pregnant people and an additional 52,884 instances of OGTT were included. Thresholds defined in the Australian pandemic guideline appear adequate to identify most GDM cases; false negative cases appeared at lower risk of hyperglycaemia in pregnancy-related events. For UK and Canadian guidelines, a larger proportion would be misdiagnosed as non-GDM; these false negative cases had broadly equivalent HIP-related event rates as true positives.

Conclusions: The OGTT remains the most effective test to identify abnormal glucose processing in pregnancy, supporting the prompt return to standard guidelines post-pandemic. Cohort studies investigating the impact of the change in guidelines on GDM pregnancies and associated outcomes are needed.

© 2021 The Author(s). Published by Elsevier Masson SAS. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Keywords: Gestational diabetes mellitus COVID-19 SARS-COV-2 Hyperglycaemia in pregnancy Scoping review

Introduction

Gestational diabetes mellitus (GDM) is a common condition [1], defined as any degree of glucose intolerance commencing in pregnancy [2]. It is of particular concern due to the adverse outcomes associated with increased glucose in pregnancy, including macrosomia, neonatal hypoglycaemia, and stillbirth [3]. In addition, people with GDM are at increased risk of GDM in future pregnancies and the development of type 2 diabetes mellitus [4]. Treatment of GDM reduces the risk of perinatal morbidity [5].

The current gold standard test is the Oral Glucose Tolerance Test (OGTT), typically involving an overnight fast, a fasting plasma glucose (FPG) test, administration of a 75 g glucose challenge, and a subsequent plasma glucose sample after 2 h [6]. As a result of the Hyperglycaemia and Adverse Pregnancy Outcomes study [3], an international observational study of 25,505 pregnant people, the International Association of Diabetes and Pregnancy Study Groups (IADPSG) and the World Health Organisation (WHO) [7] recommended universal screening of pregnant people between 20 and 24 weeks using a 75 g-OGTT, with cut-off thresholds of FPG 5.1 - 6.9 mmol/L, 1-hour plasma glucose (PG) >10 mmol/L, and 2-hour PG 8.5 – 11 mmol/L. However, some have raised concerns about increasing medicalisation of mild hyperglycaemia in pregnancy with minimal effect on adverse outcomes, and the lack of randomised controlled trials to support IADPSG’s thresholds [8]. Alternatives to the OGTT, such as continuous glucose monitoring or the use of biomarkers, have also been proposed due to concerns about test reproducibility and performance challenges [9]. Many countries, such as the United Kingdom, have opted for different threshold values or for selective as opposed to universal screening.

The emergence of the novel coronavirus, SARS-COV-2, in December 2019 has highlighted the challenge of selecting optimal tests and
thresholds. As health services prepared to care for patients with severe COVID-19, and those at risk of severe disease – including pregnant people – were advised to isolate from society, obstetrics Colleges and institutions across the globe made recommendations to limit time spent in healthcare settings and the pressures on antenatal services. New guidance was issued, replacing the OGTT with various thresholds for FPG, RPG, and HbA1c. However, while a number of systematic reviews and large-scale observational studies have previously examined FPG [10,11], random plasma glucose (RPG) [12,13], and glycated haemoglobin (HbA1c) [14,15], there remains a lack of consensus around the optimal sensitivity and specificity of the thresholds recommended in the COVID-19 pandemic GDM testing guidelines. High sensitivity might typically be preferred due to the adverse outcomes associated with not diagnosing and treating GDM in those with abnormal glucose tolerance, but in a pandemic low specificity would result in high numbers of false positives, which puts pressure on overwhelmed health services [16].

This clinical and health service dilemma, posed by the balance in risks (of COVID-19 transmission) versus benefit (in accurate detection of GDM) is evolving and emerging alongside the evolution of the COVID-19 pandemic itself. We identified a scoping review as an appropriate methodology for determining the scope and coverage of the peer-reviewed scientific literature investigating the testing strategies proposed and implemented for GDM during the pandemic, compared to the pre-pandemic diagnostic testing guidelines [17].

Methods

We performed a scoping review of the available international literature reporting studies investigating the performance of diagnostic tests for GDM that have been recommended for use during the COVID-19 pandemic, and that have evaluated these against pre-pandemic reference standard tests. We also sought to investigate the potential test and clinical outcomes for pregnant people resulting from the change in diagnostic criteria.

Literature search

A search was conducted on four databases Embase (OvidSP) [1946-present], Lit Cvid (https://www.ncbi.nlm.nih.gov/research/coronavirus/), Cochrane COVID-19 Study Register (https://covid-19.cochrane.org/) and medRxiv (https://mcguinlu.shinyapps.io/medr xiv/) on 14th June 2021. The search strategy was developed in conjunction with a Medical Information Specialist at the Bodleian Health Care Libraries, Oxford. The electronic search included the key terms (COVID-19 OR coronavirus) AND (gestational OR pregnancy OR pregnant) AND (diabetes OR hyperglycaemia OR hyperglycaemia). No date, language, or publication type limits were applied to the search. The 227 results were exported to Endnote X9 where 75 were excluded as duplicates. The remaining 145 papers were independently screened by title and abstract by two reviewers (AMC and AJF) using Rayyan software; for any disagreement, a third reviewer (AJF) made a final decision.

Inclusion and exclusion criteria

The following criteria were required for inclusion in this scoping review:

- Any primary observational study, whether prospective or retrospective, cross-sectional or cohort

In order to include studies investigating GDM testing strategies recommended for use during the COVID-19 pandemic, the search was limited to studies published after December 31, 2019, as this is regarded as the date of the first notification of SARS-COV-2 to the WHO [18].

During screening, we excluded articles which did not include an eligible population (n = 67), intervention (n = 13), outcome (n = 8), publication type (n = 32), or study design (n = 12). We identified 22 full-text articles to screen for inclusion; of these, 13 met the criteria for inclusion (Fig. 1).

Data extraction

Included articles were compiled in a custom Excel data-extraction form. The extracted data contained study characteristics and outcomes. Study characteristics included author, title, publication year, country, design, setting, inclusion and exclusion criteria, study period, sample size, participant characteristics (age, BMI, and ethnicity), and primary and secondary outcomes. Outcomes included reference test, index test(s), diagnostic threshold(s), discrepancy between index and reference (N (%) of people missed using pandemic diagnostic test specifications), sensitivity of index test, referenced COVID diagnostic thresholds used, screening type (universal or selective), and gestation at which testing was performed.

Data extraction was performed by one reviewer (AMC) and each item was checked by a second reviewer (LCA).

Data synthesis

We extracted, where possible, reported accuracy parameters (true positive, false negative, true negative and false positive) at all reported thresholds for each index test in each individual study and the data from which these were derived. From these, we calculated sensitivity and specificity data for each threshold of each index test.

Where the population studied all had GDM (diagnosed by the reference standard) or there was no information provided on false positives in the data set, we were unable to extract specificity data.

Results

The 13 studies included 40,836 pregnant people and an additional 52,884 instances of OGTT. Characteristics of the studies are reported in Table 1.

Thresholds for diagnosing gestational diabetes mellitus: pre- and post-COVID

As their reference standard test, 7 studies used IADPSG criteria as their thresholds for diagnosing GDM [32]. Three used NICE criteria [33]. Three studies used multiple criteria for diagnostic thresholds; data for each threshold were extracted separately. Details of the reference standard guidelines are contained in Table 2.

The index tests examined by these studies, used during the COVID-19 pandemic, were derived from five guidelines. Five studies examined diagnostic thresholds based on Australian guidelines from the Australasian Diabetes in Pregnancy Society (ADIPS) and the Royal Australian and New Zealand College of Obstetricians and Gynaecologists (RANZ COG), one study used criteria from the Japanese Society of Diabetes and Pregnancy and three used criteria from the Royal College of Obstetricians and Gynaecologists (RCOG) in the UK. One study used the French Society for Diabetes (SFD) and French National College of Obstetricians and Gynaecologists (CNGOF) joint guideline [30]. One study used Australian, Canadian and UK pandemic guidelines in their
Table 1
Characteristics of the included studies. GDM = Gestational Diabetes Mellitus; OGTT = Oral Glucose Tolerance Test; HAPO = Hyperglycaemia and Adverse Pregnancy Outcomes.

| Primary author, year | Country | Design | Sample size | Inclusion criteria | Study period |
|----------------------|---------|--------|-------------|--------------------|--------------|
| Siru [19] (2020)     | Australia | Retrospective laboratory audit | 16,169 | Obstetrician-referred OGTTs performed by community-based lab | Jan 2017 – Apr 2020 |
| Van Gemert [20] (2020) | Australia | Retrospective laboratory audit | 1922 | Pregnant people diagnosed with GDM using OGTT | Dec 2013 – 2019 |
| Zhu [21] (2021)      | Australia | Retrospective hospital audit | 237 | Diagnosed with GDM based on OGTT weeks 1–26 | Jan 2019 – Feb 2020 |
| Van de l’Isle [22] (2020) | UK | Prospective observational | 15,734 | All pregnant people received risk-factor screening | Jan 2016 – Dec 2019 |
| Meek [23] (2020)     | UK | Retrospective audit and prospective observational study data | 18,923 | Three cohorts: 1) Consecutive singleton pregnancies with liveborn infants 2) Consecutive GDM pregnancies 3) Pregnant people with ≥1 GDM risk factor | Jan – Dec 2019 |
| Issa [24] (2020)     | UK | Retrospective hospital audit | 205 | All pregnant people attending Wythenshawe Hospital diabetes antenatal clinic in 2019 | Jan – Dec 2019 |
| Kasuga [25] (2020)   | Japan | Retrospective hospital audit | 264 | Pregnant people diagnosed with GDM receiving perinatal care at study hospital | Jan 2013 – Dec 2019 |
| McIntyre [26] (2020) | International | Retrospective observational study | 5974 | HAPO study participants with singleton pregnancies No treatment for GDM | Jan 1 – Jun 30, 2015 |
| D’Emden [27] (2020)  | Australia | Retrospective laboratory audit | 26,242 | All OGTTs performed at selected laboratories in study period | 1 Jan 2019 – 31 Dec 2015 |
| Ikomi [28] (2020)    | UK | Retrospective hospital audit | 3805 | People with OGTT 22–30 weeks, 18–50 years old, single foetus pregnancies, no history of diabetes/bariatric surgery | Jan 2012 – Oct 2016 |
| Nachtergaele [29–31] (2021) | France | Retrospective hospital audit | 4245 | | |
One study examined thresholds that corresponded not explicitly referenced. Table 3 details the guidelines developed for diagnosis of GDM during the COVID-19 pandemic.

Table 3

| Gestation | Screening details | Guidelines, by nation | Reference standard guidelines |
|-----------|-------------------|-----------------------|-------------------------------|
| 12 weeks  | Screening strategy | Universal Preferred: Universal screening with 50 g glucose challenge, then diagnostic 75 g OGTT if 1 hr glucose challenge = 7.8 – 11.0 mmol/L (41 mmol/mol) or HbA1c ≥ 5.6% (39 mmol/mol) | Diabetes Canada Clinical Practice Guidelines Expert Committee [34] |
| 28 weeks  | Screening strategy | Universal OGTT 75 g | Queensland Clinical Guideline [35] |

Table 2

Pre-COVID reference standard guidelines. OGTT = Oral Glucose Tolerance Test; FPG = fasting plasma glucose; HbA1c = glycated haemoglobin; BMI = body mass index; GDM = gestational diabetes mellitus; US = ultrasound; RPG = random plasma glucose; T2DM = Type 2 diabetes mellitus.

Gestation  | Screening details  | Reference standard guidelines  |
-----------|--------------------|--------------------------------|
          |                    | Diabetes Canada Clinical Practice Guidelines Expert Committee [34] |
          |                    | Queensland Clinical Guideline [35] |
          |                    | International Association of Diabetes and Pregnancy Study Groups [32] |
          |                    | National Institute for Health and Care Excellence [33] |

12 weeks  | Screening strategy | Perform HbA1c  |
|          | Thresholds for diagnosis | T2DM if HbA1c ≥ 5.6% (48 mmol/mol) or GDM if FPG ≥ 5.1 mmol/L (90 mg/dL) or HbA1c ≥ 5.9% (41 mmol/mol) |

28 weeks  | Screening strategy | Universal OGTT 57 g | Universal 75 g OGTT |
|          | Thresholds for diagnosis of GDM | FPG ≥ 5.1 mmol/L, or 1 h PG ≥ 10 mmol/L, or 2 h PG ≥ 8.5 mmol/L. | FPG ≥ 5.1 mmol/L, or 1 h PG ≥ 10 mmol/L, or 2 h PG ≥ 8.5 mmol/L. |

Analysis: data were extracted separately for each national guideline. One study did not explicitly detail the use of specific diagnostic criteria as recommended during the COVID-19 pandemic, but instead investigated the sensitivity and specificity of various thresholds for diagnosis; from this study, thresholds were extracted that corresponded with those in the Japanese, Canadian, Australian and UK guidelines. One study examined thresholds that corresponded with RCOG. French and Japanese FPG guidelines, although these were not explicitly referenced. Table 3 details the guidelines developed for diagnosis of GDM during the COVID-19 pandemic.

Index tests used for gestational diabetes mellitus diagnosis

All of the COVID-19 GDM-testing guidelines recommended the use of FPG to diagnose GDM, in place of OGTT recommended pre-pandemic; some also recommend diagnostic thresholds for RPG or HbA1c. Of the included papers, all tested FPG as an index test against OGTT as the reference gold standard (OGTT with IADPSG/WHO criteria or NICE criteria). Three also tested RPG, and five tested HbA1c.

Table 3

Guidelines for diagnosis of GDM proposed during the COVID-19 pandemic. OGTT = Oral Glucose Tolerance Test; FPG = fasting plasma glucose; HbA1c = glycated haemoglobin; LGA = large-for-gestational-age; GDM = gestational diabetes mellitus; US = ultrasound; RPG = random plasma glucose; T2DM = Type 2 diabetes mellitus.
Sensitivity and specificity of recommended random plasma glucose thresholds

RCOG and JDS recommended a diagnostic threshold for detecting GDM using RPG at 28 weeks gestation of ≥9 mmol/L, while Diabetes Canada recommended a threshold of ≥11.1 mmol/L at 28 weeks. Of the papers that provided data on RPG, all demonstrated low sensitivities (0% [25], 30% [23]) at a threshold of ≥9.0 mmol/L at 13–26 weeks and 12 weeks respectively (Table 4). Specificity data for the threshold was 98% at 12 weeks [23]. One paper applied RCOG criteria but did not provide detail of the proportion of people with an RPG over the threshold [22].

We identified no studies evaluating the Diabetes Canada RPG threshold of ≥11.1 mmol/L for diagnosing GDM; the highest threshold tested was ≥10.0 mmol/L, which had a sensitivity of 11–12% at 24–28 weeks gestation [23]. Specificity for RPG at 12 weeks at thresholds from 7.0 to 10.0 mmol/L, however, was high [23]. No papers provided specificity data for RPG at 28 weeks.

Sensitivity and specificity of recommended glycated haemoglobin thresholds

For HbA1c, most guidelines (RCOG; Diabetes Canada; JDS; SFD and CNGOF) have recommended a threshold of ≥5.7% (39 mmol/mol) for diagnosis of GDM during the COVID-19 pandemic. This threshold had poor sensitivity for the diagnosis of GDM, with a range of 13–26% across the studies testing this threshold (Table 4). Only one paper provided specificity (97%) at this threshold [23].

In Australia, a HbA1c has been recommended in the first trimester for people with risk factors for GDM; Zhu et al. [21] tested this strategy using a threshold of HbA1c ≥5.9% (41 mmol/mol) and found that this would miss 98% of GDM cases, with a specificity of 2%. In their sample of 237 pregnant people with IADPSG-diagnosed GDM, none had a 1st-trimester HbA1c ≥5.9% (41 mmol/mol) and only 5% (2%) had a 2nd-trimester HbA1c ≥5.9% (41 mmol/mol). Meek et al. tested an HbA1c threshold of ≥5.8% (40 mmol/mol) in the second trimester, which had a sensitivity of 15–17% [23].

Sensitivity and specificity of recommended fasting plasma glucose thresholds

The Australian guidelines recommend people with FPG ≤4.6 mmol/L receive no further investigation or treatment, whilst those with a FPG 4.7–5.0 mmol/L should receive an OGTT as per IADPSG criteria, and those with ≥5.1 mmol/L should be diagnosed with GDM [36]. A threshold of ≥4.7 mmol/L was applied by multiple studies and yielded a good sensitivity ranging between 54 – 75% (Table 4). One paper provided specificity (100%) at this threshold [29].

The RCOG guidelines recommend FPG ≥5.6 mmol/L – or ≥5.3 mmol/L depending on resources, clinical capacity, or population characteristics – are diagnosed as GDM during the COVID-19 pandemic [40]. People below these thresholds receive no further treatment. A threshold of 5.6 mmol/L was applied in 3 separate studies and sensitivities ranged from 13 to 36%, with specificity of 100% (Table 4).

Alongside sensitivity, some studies presented the likely proportion of patients with GDM who would be missed according to the pandemic diagnostic criteria. One study prospectively tested RCOG criteria (with an FPG ≥5.3 mmol/L) on high-risk pregnant people during the pandemic and found that this missed 59% of cases that were later diagnosed with an OGTT or other glucose monitoring according to NICE criteria [22]. Another study tested RCOG criteria (FPG ≥5.6 mmol/L or HbA1c ≥5.7% [39 mmol/mol]) and found that 81% of

Table 4

| Study | Gold standard reference | Sample size (n with GDM) | Gestation at testing (weeks) | Index test | Index threshold | Sensitivity (%) | Specificity (%)# |
|-------|------------------------|-------------------------|-----------------------------|------------|----------------|----------------|------------------|
| Zhu   | IADPSG                 | 237                     | 1–26                        | FPG        | ≥4.7 mmol/L*   | 75             | –                |
|       |                        |                         |                              | HbA1c      | ≥5.9% (41 mmol/mol) | 2             | –                |
| Van Gemert | IADPSG              | 1992                    | 1–32                        | FPG        | ≥4.7 mmol/L*   | 71             | –                |
| McIntyre | IADPSG               | 1014                    | 24–32                       | FPG        | ≥4.7 mmol/L*   | 75             | –                |
| Diabetes Canada | IADPSG       | 555                     |                              | HbA1c      | ≥5.7% (39 mmol/mol) | 18            | –                |
| Siru | IADPSG                | 1790                    | 24–28                       | FPG        | ≥4.7 mmol/L*   | 54             | –                |
| D’Emden | IADPSG               | 3946                    | 24–32                       | FPG        | ≥4.7 mmol/L*   | 73             | –                |
| Kasuga | IADPSG               | 264                     | 13–26                       | FPG        | ≥5.1 mmol/L    | 34             | –                |
| Nachtergaele [30] | IADPSG   | 467                     | 22–30                       | FPG        | ≥5.0 mmol/L    | 13             | –                |
| Nachtergaele [31] | NICE      | 312                     | 22–30                       | HbA1c      | ≥5.7% (39 mmol/mol) | 15            | –                |
| Nachtergaele [29] | IADPSG   | 481                     | 22–30                       | FPG        | ≥5.6 mmol/L*   | 33*           | –                |
| Ikomi | IADPSG                | 694                     | 24–28                       | FPG        | ≥4.7 mmol/L*   | 69*           | 100              |
| Issa | NICE                  | 152                     | 28                           | FPG        | ≥5.1 mmol/L    | 43             | –                |
|       |                        |                         |                              | HbA1c      | ≥5.7% (39 mmol/mol) | 12            | –                |
| Meek | IADPSG****             | 776                     | 24–32                       | FPG        | ≥5.3 mmol/L    | 50             | –                |
|       |                        |                         |                              | HbA1c      | ≥5.7% (39 mmol/mol) | 12            | 97               |
|       |                        |                         |                              | FPG        | ≥5.6 mmol/L    | 13             | 100              |
|       |                        |                         |                              | HbA1c      | ≥5.7% (39 mmol/mol) | 30            | 97               |
|       |                        |                         |                              | FPG        | ≥5.6 mmol/L    | 25             | 100              |
|       |                        |                         |                              | HbA1c      | ≥5.7% (39 mmol/mol) | 40            | 100              |
|       |                        |                         |                              | FPG        | ≥5.0 mmol/L    | 30             | 98               |
|       |                        |                         |                              | HbA1c      | ≥5.7% (39 mmol/mol) | 41*           | –                |
|       |                        |                         |                              | RPG        | ≥5.3 mmol/L    | 41*           | –                |
|       |                        |                         |                              | HbA1c      | ≥5.7% (39 mmol/mol) | 41*           | –                |
|       |                        |                         |                              | RPG        | ≥9.0 mmol/L    | –              | –                |

*ADIPS / RANZCOG guidelines state that pregnant people with FPG ≥4.7 mmol/L are referred for further OGTT if between 4.7–5.0 mmol/L and diagnosed as GDM if ≥5.1 mmol/L. All pregnant people with an FPG <4.7 mmol/L would therefore be classified incorrectly as non-GDM. *Value provided is the sensitivity for the combination of thresholds; individual proportions were unavailable. **IADPSG data solely shown here due to complexity; NICE data can be found in supplementary appendix. *Data presented is for universal screening only; selective screening data found in supplementary appendix. *Specificity data unavailable for some papers.
people who would previously have been diagnosed under NICE would be classified as needing no further treatment [26]. The Japanese and French guidelines recommend FPG ≥5.1 mmol/L are diagnosed as GDM, and those with results below this threshold be classified as needing no further treatment. Five studies tested this proposal and found sensitivities ranging from 34 – 64% [23,25,28–30] and specificity of 100% [23,28,29].

### Hyperglycaemia in pregnancy-related events

Five studies provided data on HIP-related events and their relationship with COVID-19 GDM guidelines [23,26,29–31]. In separate analyses using the RCOG and SFD pandemic criteria, Nachtergaele et al. found no difference in rates of HIP-related events in the false negative group versus the true positive group [30,31]. However, the true positive group was significantly more likely to require insulin therapy during pregnancy. Another analysis exploring a variety of screening criteria – most similar to the Australian pandemic criteria – found that false negative cases had decreased rates of some HIP-related events, but not all, compared to true positives [29].

Meek et al. noted that the positive index tests (RPG at 12 weeks, FPG and HbA1c at 28 weeks) had associations with some, but not all, of the pregnancy outcomes of LGA infant, neonatal hypoglycaemia, admission to NICU, and Caesarean section [23].

Using the HAPO study population, McIntyre et al. observed that false negative cases of GDM, who were misclassified by Canadian and UK pandemic criteria, had statistically significantly higher rates of HIP-related pregnancy events compared to the non-GDM (as defined by reference standard) population. Rates of primary Caesarean section differed between false negatives by pandemic criteria (Canadian: 25%; UK: 24%), true positives (Canadian: 25%; UK: 22%), and those non-GDM by both pre- and post-COVID criteria (Canadian: 18%; UK: 17%). Rates of neonatal hypoglycaemia also differed between false negatives by pandemic criteria (Canadian: 21%; UK: 19%), true positives (Canadian: 26%; UK: 27%), and non-GDM by both pre- and post-COVID criteria (Canadian: 17%; UK: 17%) [26].

### Discussion

#### Principal findings

We aimed to scope the existing body of evidence on the impact of temporary guidelines for GDM diagnoses during the COVID-19 pandemic on the screening and diagnosis of GDM internationally, alongside assessing the sensitivity and specificity of the guideline thresholds used and capturing the rate of hyperglycaemia-in-pregnancy (HIP)-related events in pre-COVID-19 GDM populations and post-COVID-19 GDM populations, the latter being identified using guidelines adapted for diagnosis of GDM during the COVID-19 pandemic. We identified thirteen studies that explored five national pandemic guidelines and a variety of thresholds for the index tests FPG, RPG and HbA1c. No single test or combination of tests identified the same rate of GDM cases as the pre-pandemic reference standards that used OGTT.

All studies used FPG as an index test. GDM was diagnosed (or an OGTT performed) at thresholds above 4.7 mmol/L (sensitivity 54–75%; specificity 100%) [19–21,26,27,29], 5.1 mmol/L (sensitivity 34–64%; specificity 100%) [23,25,28–30], 5.3 mmol/L (sensitivity 25–50%; specificity 100%) [22–24], and 5.6 mmol/L (sensitivity 13–43%; specificity 100%) [23,24,28,31].

Most studies also compared HbA1c against the reference standard. GDM was diagnosed at thresholds above 5.7% (39 mmol/mol) (sensitivity 12–30%; specificity 97%) [22–26,30,31] or above 5.9% (41 mmol/mol) (sensitivity 2%; specificity unavailable) [21].

Few studies explored RPG as a diagnostic tool for GDM, despite its inclusion in three national guidelines. GDM was diagnosed at thresholds above 9.0 mmol/L (sensitivity 0–30%; specificity 98%) [22,23,25].

Five studies explored the impact of altered guidelines on the identification of HIP-related pregnancy and neonatal events.

#### Previous literature

Previous studies have challenged the idea of uniform worldwide diagnostic thresholds, such as those recommended by the WHO/ IADPSG. Practic pragmatic adoption of index test thresholds in the absence of OGTT, as a result of the COVID-19 pandemic, has led to new evaluations based on these temporary guidelines and potential alternative pathways and thresholds to be considered going forward.

Various systematic reviews have explored individual index tests and their thresholds for the diagnosis of GDM. A previous systematic review of FPG as a GDM diagnostic in 2013 found that a cut-off of 5.1 mmol/L had a sensitivity of 68.6% and specificity of 93.2% — broadly similar to the findings in this review [10].

Another systematic review with a cut-off of 4.7 mmol/L found sensitivity of 87% and specificity of 52% [11]. A systematic review and meta-analysis of HbA1c in GDM indicated that a threshold of 5.7% (39 mmol/mol) had high specificity (90%), but poor sensitivity (36%). The optimal threshold for sensitivity was 5.2% (33 mmol/mol), but this had low specificity and would result in many false positives. Another review found sensitivity and specificity of 25% and 96% respectively at a cut-off of 5.7% (39 mmol/mol); higher thresholds had higher specificity and lower sensitivity. Pillay et al. found that HbA1c at any cut-off did not achieve both high sensitivity and specificity [5].

A systematic review of RPG testing concluded that RPG was inadequate to diagnose GDM, with low sensitivity and specificity across a range of thresholds. A large study of RPG taken at 12–16 weeks in 17,736 pregnant people found that a cut-off of ≥7.5 mmol/L had sensitivity of 70% and specificity of 90% for the detection of GDM; RPG was also a better predictor of GDM than maternal age or BMI. A threshold of ≥8.5 mmol/L had sensitivity at 43% and specificity at 97% compared against IADPSG; a cut-off ≥4.7 mmol/L was required for a sensitivity of 90% [13].

For RPG, HbA1c, and RPG, higher sensitivity often means compromising on specificity, and vice versa. In non-pandemic times, a higher sensitivity would typically be preferred as this reduces the likelihood of misdiagnosing GDM and failing to treat pregnant people with abnormal glucose metabolism in pregnancy. Evidence from the HAPO study suggests that increasing glucose is correlated with increasing likelihood of HIP-related events. However, higher sensitivity requires lower index test thresholds and risks medicalising pregnancies that may not be at increased risk of HIP-related events. The current best available evidence is observational, with no randomised controlled trials demonstrating risk of harm of mild hyperglycaemia in pregnancy. McIntyre et al. [26] found that participants with FPG below 4.7 mmol/L had lower rates of pregnancy complications than those with FPG above this threshold, and in the HAPO study the subgroup with the greatest odds ratios for LGA infant had an elevated FPG and elevated post-load plasma glucose [43].

### Strengths and limitations at study and outcome level

The included studies are all performed in high-income countries. Despite no language restrictions for our searches, we found no studies in middle- or low-income countries. There may be ethnic differences that limit the generalisability of some of the findings; for example, a Japanese population may have decreased insulin secretion compared to a majority-Caucasian population [25].

Interpreting HIP-related event data also has notable limitations. Four of the studies were retrospective analyses, where participants...
diagnosed by the reference standard were treated for GDM; there-fore, any conclusions drawn about likely clinical outcomes for people classified as false negative cases may be inaccurate, as they received treatment for GDM. However, McIntyre et al. used a HAPO study pop-ulation for their retrospective analysis [26]. Given that no HAPO par-ticipants with a FPG of 5.8 mmol/L or less, or 2 h plasma glucose of 11.1 mmol/L or less, received treatment, these findings may be more in line with real-world outcomes for missed cases of mild hypergly-cæmia.

Many of the studies included in this review are retrospective analy-ses of pre-COVID populations, with COVID-adapted diagnostic thresholds applied. Few looked prospectively at the impact on preg-nant people of real-time change in diagnostics; however, given the risk of COVID-transmission in healthcare settings situation, it may have been inappropriate to put patients at risk of infection by follow-ing pre-COVID GDM guidelines. Van-de-Pisle et al. noted that, when retested with an OGTT, pregnant people in their prospective sample had a higher screen-positive rate for GDM [22]. They hypothesise that the UK pandemic lockdown may have resulted in decreased daily exercise and increased anxiety, both of which are associated with increased risk of developing GDM [44]. Future studies will ide-ally evaluate the direct and clinical impact of alternative testing stra-tegies during COVID-19; however, it is worth noting that these populations will be impacted by the unprecedented pandemic, and rates of GDM and associated HIP-related events may differ signifi-cantly as a result. While retrospective analyses have some limitations, they do allow direct comparison of guidelines within a population.

Strengths and limitations at review level

We performed a comprehensive scoping search with a systematic approach, giving methodological rigour to this scoping review. We included primary observational studies in all languages, including grey literature. In addition, this review is the first attempt to scope the existing body of literature on the different COVID-19 GDM screening and diagnostic criteria and their impact on the identifica-tion of GDM and resultant HIP-related events; as such, it has real-world application internationally.

Though we did not assess the quality of studies using a structured framework, as in a systematic review, the societal and clinical cir-cumstances likely constrained study design for this clinical issue and may have impacted research quality. We have identified sufficient studies such that a full systematic review could be performed; this could include a formal assessment of study quality.

Clinical implications

As noted by RCOG, diagnostic strategies adapted for the COVID-19 pandemic are only suitable for use at the peaks of the pandemic and there is an expectation that clinicians should return to normal diag-nostic guidelines as soon as prevalence of SARS-COV-2 and local trans-mission allows [40]. The findings of the studies captured by this review support a return to the gold standard.

However, FPG, RPG and HbaA1c do have some potential to identify and stratify patients with GDM: at lower thresholds, FPG can ade-quatley identify people at risk of raised post-load plasma glucose. Some health services may in future consider using a two-step model of FPG and subsequent OGTT, similar to the ADIPS and RANZCOG pan-demic guidelines [36,37]. Previous reviews have also highlighted the importance of HbaA1c or FPG testing in early pregnancy to identify pre-pregnancy abnormal glucose processing and intervene appropri-ately to mitigate HIP-related events [9].

Some of the studies in this review evaluated selective screening, where risk factors are used to identify pregnant people at increased likelihood of developing GDM. Previous studies have explored the poorer sensitivity of selective screening for GDM detection [45]; post-pandemic, services could look to replace or add to risk-factor-based screening with a universal FPG. A model combining FPG with a risk prediction score has demonstrated good accuracy for detection of GDM [46].

At higher thresholds, FPG, RPG and HbaA1c misclassify more cases of GDM, including those at risk of HIP-related events. For services that have used these thresholds, identifying those with undiagnosed GDM will be essential, as these individuals have a 50% risk of Type 2 diabetes mellitus within 5 years of their delivery and a 70% risk of GDM in a future pregnancy [47,48], as well as potential long-term cardiometabolic consequences for children [9].

Conclusions

This review scoped a variety of GDM guidelines introduced during the pandemic and investigated the sensitivity and specificity of index tests recommended in these guidelines, against reference-standard tests for GDM using the identified relevant scientific studies. Guide-lines adapted for the COVID-19 pandemic have been shown to consist-ently miss a proportion of patients with GDM when applied to retrospective datasets. Future research should look to prospectively evaluate the diagnostic performance of these tests. Furthermore, it will be important to explore the rate of HIP-related events associated with GDM in patients who have delivered during the COVID-19 pan-demic, comparing rates amongst those who were diagnosed with GDM versus those who were not, and seeking to understand whether comparative rates differ to those when the gold-standard guidelines have been in use.

Contributions

AMC, LCA and AJF conceived of and designed this scoping review. AMC, LCA, NW, and AJF developed the search strategies. NW per-formed database searches. AMC and LCA performed citation screen-ing and data extraction for the included studies. AMC led the writing of the manuscript and all authors contributed to successive drafts, approved the final manuscript, and agree to be accountable for all aspects of the work.

Declaration of Competing Interest

AJF is Director of the NIHR Health Technology Assessment Pro-gramme. The authors declare that there are no known competing financial interests or personal relationships that could influence or appear to influence the work reported in this paper.

Funding

No funding was received to perform this scoping review. This research was supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. LCA is funded by a Wellcome Trust Doctoral Fellowship [203921/Z/16/Z]. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.demana.2021.100023.
