



Progression to type 2 diabetes in women with a known history of gestational diabetes: systematic review and meta-analysis

Elpida Vounzoulaki,^{1,2} Kamlesh Khunti,^{1,2} Sophia C Abner,^{1,2} Bee K Tan,³ Melanie J Davies,¹ Clare L Gillies^{1,2}

¹Diabetes Research Centre, Leicester General Hospital, University of Leicester, Leicester, UK

²National Institute for Health Research Applied Research Collaboration-East Midlands, Leicester Diabetes Centre, Leicester LE5 4PW, UK

³Department of Cardiovascular Sciences, University of Leicester, Leicester, UK

Correspondence to: E Vounzoulaki

ev63@le.ac.uk

(or @Elpida_Vnz on Twitter ORCID 0000-0003-3137-8295)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2020;369:m1361 <http://dx.doi.org/10.1136/bmj.m1361>

Accepted: 24 March 2020

ABSTRACT

OBJECTIVE

To estimate and compare progression rates to type 2 diabetes mellitus (T2DM) in women with gestational diabetes mellitus (GDM) and healthy controls.

DESIGN

Systematic review and meta-analysis.

DATA SOURCES

Medline and Embase between January 2000 and December 2019, studies published in English and conducted on humans.

ELIGIBILITY CRITERIA FOR SELECTING STUDIES

Observational studies investigating progression to T2DM. Inclusion criteria were postpartum follow-up for at least 12 months, incident physician based diagnosis of diabetes, T2DM reported as a separate outcome rather than combined with impaired fasting glucose or impaired glucose tolerance, and studies with both a group of patients with GDM and a control group.

RESULTS

This meta-analysis of 20 studies assessed a total of 1 332 373 individuals (67 956 women with GDM and 1 264 417 controls). Data were pooled by random effects meta-analysis models, and heterogeneity was assessed by use of the I^2 statistic. The pooled relative risk for the incidence of T2DM between participants with GDM and controls was estimated. Reasons for heterogeneity between studies were investigated by prespecified subgroup and meta-regression analyses. Publication bias was assessed by funnel plots and, overall, studies were deemed to have a low risk of bias ($P=0.58$ and $P=0.90$). The overall relative risk

for T2DM was almost 10 times higher in women with previous GDM than in healthy controls (9.51, 95% confidence interval 7.14 to 12.67, $P<0.001$). In populations of women with previous GDM, the cumulative incidence of T2DM was 16.46% (95% confidence interval 16.16% to 16.77%) in women of mixed ethnicity, 15.58% (13.30% to 17.86%) in a predominantly non-white population, and 9.91% (9.39% to 10.42%) in a white population. These differences were not statistically significant between subgroups (white v mixed populations, $P=0.26$; white v non-white populations, $P=0.54$). Meta-regression analyses showed that the study effect size was not significantly associated with mean study age, body mass index, publication year, and length of follow-up.

CONCLUSIONS

Women with a history of GDM appear to have a nearly 10-fold higher risk of developing T2DM than those with a normoglycaemic pregnancy. The magnitude of this risk highlights the importance of intervening to prevent the onset of T2DM, particularly in the early years after pregnancy.

SYSTEMATIC REVIEW REGISTRATION

PROSPERO CRD42019123079.

Introduction

Gestational diabetes mellitus (GDM) is glucose intolerance with onset or first diagnosis during the second or third trimester of pregnancy, which is clearly not either pre-existing type 1 or type 2 diabetes mellitus (T2DM).¹ A previous diagnosis of GDM is an established risk factor for developing T2DM in later life,² a fact that highlights the importance of postpartum screening to identify those at higher risk of progression and introduce strategies for disease prevention. Despite the magnitude of the risk, attendance for postpartum screening is suboptimal, as reported by studies from European countries, the United States, and Canada.³⁻⁶ An overall lack of awareness of the need for screening and the increased risk for future T2DM in this patient group is apparent.⁷⁻⁹ Healthcare providers lack familiarity with, and adherence to, guidelines,^{10 11} and are uncertain as to whether the responsibility for screening these women lies with primary or secondary care.¹¹ Meanwhile, the prevalence of T2DM is rising globally, with individuals diagnosed with the disease being at higher risk of all cause mortality.²

Previous systematic reviews have highlighted the increased risk of developing T2DM after GDM. Kim et al, in 2002, identified a cumulative incidence of T2DM of between 2.6% to over 70% in studies with a total follow-up ranging from six weeks to 28 years; this incidence was highest in the first five

WHAT IS ALREADY KNOWN ON THIS TOPIC

Gestational diabetes is an established risk factor for developing type 2 diabetes in later life

Previous systematic reviews published over a decade ago have investigated the risk of subsequent progression to type 2 diabetes in women with gestational diabetes

Recent high quality systematic reviews quantifying the magnitude of this risk, in ethnically diverse populations and in the long term, are lacking

WHAT THIS STUDY ADDS

This study suggests that women with a history of gestational diabetes are almost 10 times more likely to develop type 2 diabetes than healthy controls

Existing evidence on the risk of postpartum type 2 diabetes is updated by including recently published studies with longer follow-up

Promotion of postpartum screening among women with previous gestational diabetes and encouragement to adopt dietary, lifestyle, and pharmacological interventions to prevent or delay the onset of type 2 diabetes, is urgently needed

postpartum years.¹² In 2009, Bellamy et al showed a sevenfold higher risk of T2DM in patients with GDM in comparison with healthy controls (relative risk 7.43, 95% confidence interval 4.79 to 11.51).¹³ A further systematic review evaluated the rate of compliance with screening, and the prevalence of T2DM in Asian women, reporting incidences of between 2.8% and 58% in women with previous GDM.¹⁴

Over recent decades, the demography of pregnant women has changed, with an increase in the rate of women giving birth at a more advanced age, and obesity rates also rising. Both these factors have led to a rise in the prevalence of GDM, establishing the disease as an imminent concern globally.¹⁵ From a public health perspective, this increased prevalence of GDM could contribute to the rising global health burden of obesity and T2DM. To deal with this problem, research has focused on improving the quality and effectiveness of screening and diagnosis of GDM.¹⁶ In addition to changes in the guidelines for GDM, changes have also occurred in the diagnostic criteria for T2DM, in an effort to identify populations at a higher risk.¹⁷ Thus there is a need to evaluate more recent evidence on the risk of progression to T2DM in women with previous GDM compared with those with a normoglycaemic pregnancy, investigating outcomes in ethnically diverse populations, and over a longer follow-up period. This systematic review and meta-analysis aims to investigate progression rates, and factors that could determine progression to T2DM, in women diagnosed with GDM compared with those with a normoglycaemic pregnancy, using recent evidence and including an assessment of both length of follow-up and ethnicity.

Methods

This study was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)¹⁸ and Meta-analyses Of Observational Studies in Epidemiology (MOOSE) guidelines¹⁹ (supplementary material). The protocol for this systematic review and meta-analysis is registered at https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=123079.

Search strategy and inclusion criteria

We searched Ovid Medline and Embase²⁰ for observational studies published between January 2000 and December 2019, choosing 2000 as the start date, after release of the 1999 World Health Organization recommendations for performance of the 75 g oral glucose tolerance test at six weeks and onward after labour.²¹ A search strategy was developed with the assistance of a clinical librarian. The strategy included both medical subject headings and free text words covering “gestational diabetes” and “type 2 diabetes” and was restricted to studies published in English and conducted on humans. The full search strategy is provided in the supplementary material. All duplicate records were removed.

Two independent authors (EV and SCA) reviewed the titles and abstracts to identify any relevant studies. Full

text reports were then obtained and screened in detail against the following predefined eligibility/inclusion criteria: studies reporting postpartum follow-up for at least 12 months, incident physician based GDM and T2DM diagnosis reported, T2DM incidence reported separately rather than combined with impaired fasting glucose or impaired glucose tolerance, and studies with both a group of patients with GDM and a control group. Cross sectional studies and studies reporting postpartum screening attendance but not T2DM progression were excluded. All reference lists from relevant systematic reviews were hand searched for additional eligible studies.

All conference proceedings, guidelines, dissertations, commentaries, and letters were excluded. Any disagreement between the study authors was resolved by discussion or by third party consultation, until consensus was reached. The literature review and study selection process are presented in a PRISMA flowchart. Study authors were contacted to provide the full text when it was not available, and also when information needed for inclusion in the analyses was missing.

Risk of bias and study quality

We used the Newcastle-Ottawa scale, a scale proposed by Wells et al²² and designed to evaluate the quality of non-randomised studies. The version of the scale used for cohort studies consists of three categories: selection, comparability, and outcome. Based on the guidelines of the scale, a cohort study can be awarded a maximum of one star for each numbered item from the selection and outcome categories and a maximum of two stars for the category of comparability.²² A study can be awarded from zero up to nine stars.²² As the Newcastle-Ottawa scale does not adequately assess potential confounders in study analyses, further information was extracted on which confounders had been considered by each study. Publication bias was assessed with funnel plots for asymmetry using the Begg’s and Egger’s tests.²⁰

Data extraction and statistical analysis

Two authors (EV and SCA) independently undertook data extraction according to The Cochrane Handbook guidelines,²⁰ and findings were reported according to PRISMA¹⁸ and MOOSE guidance.¹⁹ Any disagreement was settled by consensus among all authors. When more than one study investigated outcomes from the same cohort, the study with information most relevant to the analysis was chosen for inclusion. Ethnic origin was defined as white, non-white, or mixed. Studies reporting a homogeneous population comprising at least 80% of the same ethnicity were categorised as white or non-white, and studies reporting a population comprising a number of different ethnicities were categorised as mixed. If ethnicity was not reported, it was defined on the basis of the predominant ethnicity of the country in which the study was conducted. When progression to T2DM was reported at a number of time points, the results from the longest follow-up point were used in the analysis.

To pool study results, random effects meta-analysis models were fitted in Stata 15.1 using the DerSimonian and Laird method,²³ study standard errors were estimated for the log transformed relative risk, and study results were pooled on the log scale.²⁴ Heterogeneity between the included studies was assessed with the I^2 statistic.²⁰ In studies in which T2DM did not occur in any of the groups, we applied a continuity correction using the default value of 0.5 to avoid any division by zero.²⁰ Any additional parameters were assessed by prespecified subgroup (categorical variables) or meta-regression (continuous variables) analysis.

To further explore the effect of ethnicity and length of follow-up, the risk of T2DM in women with and without GDM was separated, for each study; results were then pooled using random effects meta-analysis models. Subgroup analyses by study level ethnicity, length of study follow-up, study design, and screening method used in pregnancy to diagnose GDM (one step v two step) were undertaken to investigate sources of heterogeneity between studies. Meta-regression models were also fitted to estimate the effect of study heterogeneity and investigated the cumulative risk of T2DM by age, body mass index, publication year, and length of follow-up. A sensitivity analysis was undertaken to explore the effect of each individual study on the overall pooled estimate.

Patient and public involvement

Patient and public involvement sessions involving women affected by GDM are run regularly at the Diabetes Research Centre to inform our research. In addition, two of the study authors (KK and BKT) have been actively involved with primary care and women's health for many years and maintain strong community links. Both authors assisted in decisions on outcomes to be considered, protocol development, approach to analysis, and interpretation of the results, to ensure consideration was given to patient and public perspective. As this study was a systematic review and meta-analysis, patient recruitment or use of patient data was not required.

Results

Identified studies

Our initial searches yielded 7448 studies, of which 287 were selected for further evaluation using the full text (fig 1). Of these, 119 were conference proceedings, 72 did not explore outcomes in a group with GDM and a healthy control group, 24 reported postpartum follow-up for a period shorter than 12 months, 15 were not relevant to the research question, 6 did not report the incidence of T2DM, 10 were cross sectional studies, 8 included patients with a self-reported diagnosis of diabetes (either GDM or T2DM or both), 5 were letters and commentaries, 2 were systematic reviews, 2 reported screening in patients already diagnosed with impaired glucose tolerance after GDM, 1 did not compare patients with GDM with a healthy population, and 1 reported information about screening but not

progression. Tehrani et al²⁵ and Hakkarainen et al²⁶ assessed the same cohorts as Minooe et al²⁷ and Huopio et al,²⁸ respectively. We included the last two studies because these reported the most relevant data for the meta-analysis. Retnakaran et al,²⁹ Feig et al,³⁰ and Mukerji et al³¹ also assessed populations from the same database. From these three, Mukerji et al was included as this study assessed the largest cohort. One study,³² did not report the necessary data on the incidence of T2DM, and as these could not be obtained from the author, it was not included in the analysis. Finally, three studies retrieved from the reference lists of previous relevant systematic reviews were also included.³³⁻³⁵ A total of 20 studies matched all the eligibility criteria.

Study characteristics

The 20 studies included in this systematic review and meta-analysis had differing lengths of follow-up, time of screening, ethnic origin of study participants, and diagnostic criteria for both GDM and T2DM.^{27 28 31 33-49} All studies were observational, seven were described as prospective cohorts, six as retrospective cohorts, four as population based studies, two as follow-up studies, and one as a hospital based study. Most of these studies were conducted in several European countries, but studies were also from South Asia, the Middle East, the US, Australia, Canada, and Korea. Total postpartum follow-up ranged from one to 25 years. A summary of study characteristics is presented in table 1 with added definitions in box 1.

Study quality

Overall, all included studies that underwent quality assessment with the use of Newcastle-Ottawa scale received a total of six to eight stars, and were thus deemed to have a low risk of bias, using a cut-off point reported in previous reviews using the same tool.^{50 51} One study received a total of five stars, indicating high risk of bias.⁴³ Studies were adjusted for a variety of different confounders, of which the most common were maternal age, body mass index, family history of T2DM, parity, ethnicity, and socioeconomic status. A summary of the study quality assessment and a separate table showing all study confounders considered when assessing comparability are provided in supplementary tables S1 and S2.

Meta-analysis

In this meta-analysis assessing a total of 1332373 individuals (67956 women with GDM and 1264417 controls), the pooled relative risk for T2DM was almost 10 times higher in women with previous GDM than in healthy controls (relative risk 9.51, 95% confidence interval 7.14 to 12.67, $P<0.001$). Every study showed that the risk for T2DM was greater in women with GDM than in controls (fig 2).

Significant heterogeneity was seen in the overall effect estimate ($I^2=96.5\%$, $P<0.001$), and hence prespecified subgroup analyses were performed to explore potential sources. The risk for T2DM was

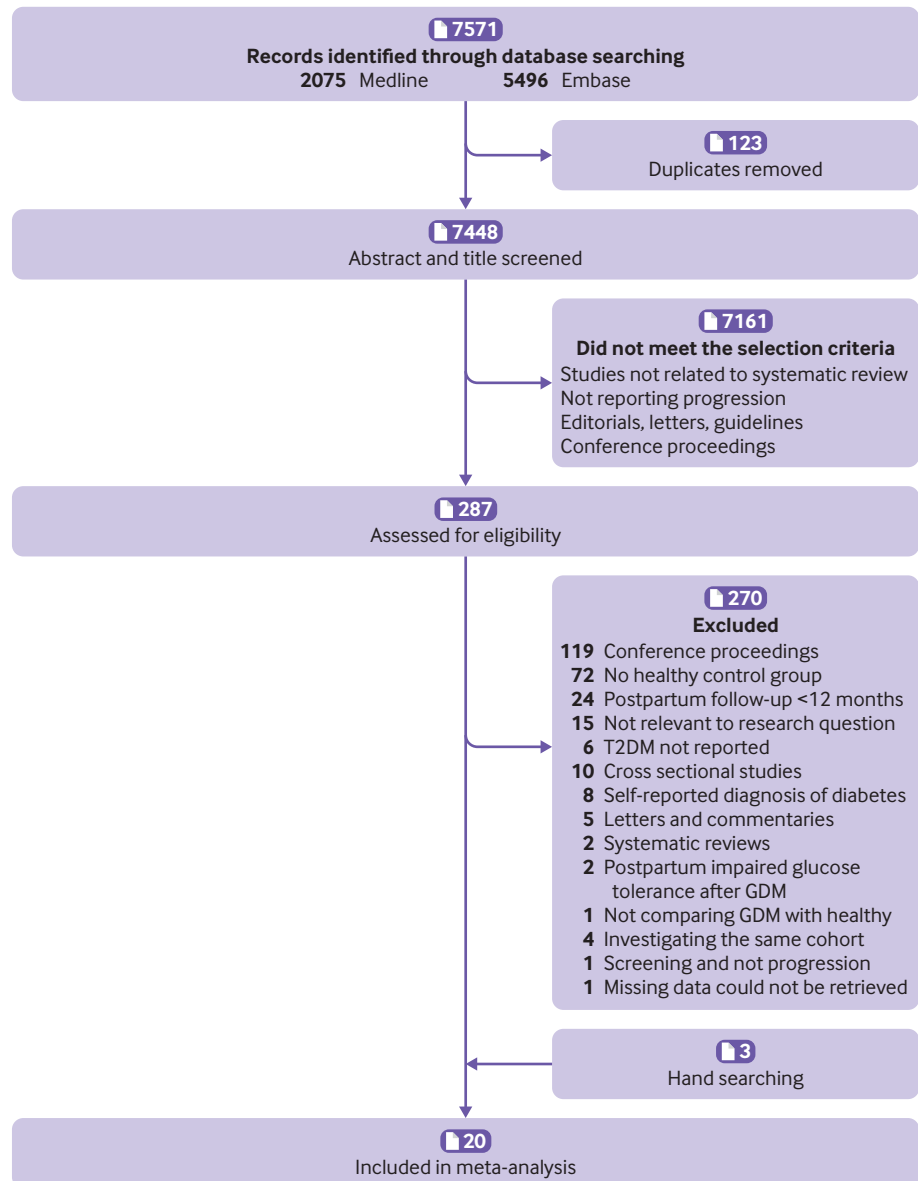


Fig 1 | Flow diagram of literature search. GDM=gestational diabetes mellitus; T2DM=type 2 diabetes mellitus

assessed by subgroup analysis for white, non-white, and mixed ethnicity as described by the studies. As shown in figure 3, the pooled relative risk for T2DM in studies with predominantly white populations was 16.28 (95% confidence interval 15.01 to 17.66), non-white 10.38 (4.61 to 23.39), and mixed populations 8.31 (5.44 to 12.69). These relative risks were not statistically significantly different between subgroups (white *v* mixed, $P=0.26$ and white *v* non-white, $P=0.54$).

The relative risk of T2DM was also assessed by subgroup analysis based on study length of follow-up. As shown in figure 4, studies were separated by their length of follow-up into three groups. The estimated pooled relative risk was 17.06 (95% confidence interval 8.95 to 32.55) for studies with follow-up of one to five years, 10.42 (5.68 to 19.11) for those with follow-up of more than five years and up to 10 years, and 8.09 (4.34 to 15.08) for those with follow-up of

more than 10 years. The differences in pooled relative risks by subgroup were not statistically significant (1-5 years *v* >5-10 years, $P=0.63$; 1-5 years *v* >10 years, $P=0.38$). Although differences were not statistically significant, the relative risk of T2DM in women with GDM compared with controls seemed to be higher in the shorter studies (<5 years; fig 4).

For a better understanding of why relative risks might have varied by ethnicity and length of follow-up, the estimated cumulative incidences for both women with GDM and controls were calculated for each study and pooled by subgroup (table 2). The cumulative incidence of T2DM was found to be higher in a population with a mix of ethnicities or a non-white population than in white populations, for both women who had had GDM, and healthy controls. These results were not statistically significant (table 2), however, and could partially be explained by the longer follow-up in studies assessing mixed ethnicities.

Table 1 | Study characteristics

Study name	Study design	Country	GDM criteria†	T2DM criteria†	Follow-up (No)	Follow-up (years)	Mean age (years)	Ethnicity
Aberg et al, 2002 ³⁶	Population based	Sweden	EASD 1991	WHO 1985	289 (GDM: 229, controls: 60)	1	Range 20-45	White
Albareda et al, 2003 ³⁵	Prospective cohort	Spain	Second and third workshop -conference on GDM	WHO 1998	766 (GDM: 696, controls:70)	11	GDM: 37.5, controls: 40	White
Aziz et al, 2018 ³⁷	Follow-up study	Pakistan	IADPSG	—	167 (GDM: 78, controls: 89)	2	GDM: 28.9, controls: 25.68	Non-white
Chodick et al, 2010 ³⁴	Retrospective cohort	Israel	Carpenter and Coustan 1982	Local	185 416 (GDM: 11 270, controls: 174 146)	5.4	GDM: 32.74, controls: 30.59	White
Daly et al, 2018 ³⁸	Retrospective cohort	UK	NICE	NICE	46 399 (GDM: 9118, controls: 37 281)	25	GDM and controls: 33.0	Mixed (white, South Asian, Afro-Caribbean, other)
Herath et al, 2017 ³⁹	Retrospective cohort	Sri Lanka	WHO 1999	WHO 1998	359 (GDM: 119, controls: 240)	10	GDM: 42.7, controls: 38.7	Non-white
Huopio et al, 2014 ²⁸	Follow-up study	Finland	Contemporary criteria	ADA 1997	874 (GDM: 489, controls: 385)	7.3	GDM: 37.8, controls: 38.4	White
Krishnaveni et al, 2007 ⁴⁰	Prospective cohort	India	Carpenter and Coustan 1982	WHO 1998	524 (GDM: 35, controls: 489)	5	GDM: 33.25, controls: 28.3	Non-white
Lee AJ et al, 2007 ⁴¹	Retrospective cohort	Australia	ADIPS 1998	WHO 1998	6253 (GDM: 5470, controls: 783)	15	GDM: 30.7, controls: 30.5	Mixed (Caucasian, Asian, Aboriginal)
Lee H et al, 2008 ^{42 *}	Prospective cohort	Korea	NDDG 1979	—	1488 (GDM: 620, controls: 868)	7	GDM and controls: 33.6	Non-white
Linné et al, 2002 ^{43 *}	Retrospective cohort	Sweden	—	WHO 1985	80 (GDM: 28, controls: 52)	15	GDM: 47.6, controls: 45.6	White
Madarász et al, 2009 ⁴⁴	Hospital based	Hungary	WHO 1985	WHO 1998	107 (GDM: 68, controls: 39)	4	GDM: 36.1, controls: 33.6	White
Minoee et al, 2017 ²⁷	Population based	Iran	WHO 1999	ADA 1997	2458 (GDM: 476, controls: 1982)	15	GDM: 36.5, controls: 34.3	Mixed (West-Asian, Eastern Mediterranean)
Mukerji et al, 2012 ³¹	Population based	Canada	—	—	1 050 108 (GDM: 33 203, controls: 1 016 905)	15	—	Mixed (Chinese, South Asian, white)
Pintaudi et al, 2015 ⁴⁵	Population based	Italy	ADA 2004	—	15 404 (GDM: 3851, controls: 11 553)	8	GDM and controls: 35.7	Mixed
Retnakaran et al, 2010 ⁴⁶	Prospective cohort	Canada	NDDG 1979	CDA 2008	180 (GDM: 107, controls: 73)	1	GDM: 35.2, controls: 35.6	Mixed (white, Asian, other)
Vambergue et al, 2008 ³³	Prospective cohort	France	Carpenter and Coustan 1982	ADA 1997	406 (GDM: 295, controls: 111)	6.8	—	Mixed (French, Maghrebian)
Vigneault et al, 2015 ⁴⁷	Prospective cohort	Canada	—	CDA 2013	299 (GDM: 216, controls: 83)	4	GDM: 36.36, controls: 35.66	Mixed (Caucasian, black, Aboriginal, Asian, Hispanic)
Wang et al, 2012 ⁴⁸	Prospective cohort	United States	Carpenter and Coustan 1982	WHO 1998/ ADA 1997	19 998 (GDM: 1142, controls: 18 856)	8.6	GDM: 26.8, controls: 24.3	Mixed (African American, white, Asian, Hispanic, Indian)
Yefet et al, 2019 ⁴⁹	Retrospective cohort	Israel	Carpenter and Coustan 1982 and NDDG 1979	—	798 (GDM: 446, controls: 352)	15.8	GDM and controls: 45.0	Mixed (Israel, Ethiopia, USSR, other)

A table of confounders considered in the analysis of each study is included in supplementary table S2.

ADA=American Diabetes Association; ADIPS=Australasian Diabetes in Pregnancy Society; CDA=Canadian Diabetes Association; EASD=European Association for the Study of Diabetes; GDM=gestational diabetes mellitus; IADPSG=International Association of Diabetes and Pregnancy Study Groups; NDDG=National Diabetes Data Group; NICE=National Institute for Health and Care Excellence; T2DM=type 2 diabetes mellitus.

*Report described as a case-control study but the methodology used indicates that it is a cohort study.

†T2DM and GDM criteria are shown in box 1.

For length of study follow-up, in women with GDM the pooled cumulative incidence was estimated to be around 9.22% (95% confidence interval 7.19% to 11.26%) in studies with one to five years of follow-up, increasing to 16.15% (15.83% to 16.47%) for studies with more than 10 years of follow-up. The cumulative incidence in controls reached 1.90% (95% confidence interval 1.87% to 1.92%) for the longest studies. The high estimated pooled relative risk for studies of less than five years of follow-up seems to be partially due to the low incidence of T2DM in the control arms of studies during this period.

Additional subgroup analyses investigated the effect of study design (prospective, retrospective, or population/hospital based; supplementary figure S1) and the screening method used in pregnancy to

diagnose GDM (one step v two step; supplementary figure S2). No significant differences were seen by study design (prospective v retrospective design, $P=0.81$; prospective v population/hospital based design, $P=0.61$) or by screening method for GDM ($P=0.58$).

Further meta-regression models found no association between study effect size and mean study age, body mass index, length of follow-up, and publication year. Results of the meta-regression analyses are presented in table 3.

In the sensitivity analysis when a named study was omitted, the pooled estimate remained close to the observed overall estimate, indicating that no individual study had a large influence on the pooled estimate. The plot for the analysis estimates is provided in supplementary figure S3.

Box 1: Type 2 diabetes mellitus and gestational diabetes mellitus criteria as used in table 1**Criteria for gestational diabetes mellitus (GDM)**

- ADA (2004)—two or more raised values during oral glucose tolerance test: fasting plasma glucose ≥ 95 mg/dL (5.3 mmol/L), one hour glucose ≥ 180 mg/dL (10.0 mmol/L), two hour glucose ≥ 155 mg/dL (8.6 mmol/L), three hour glucose ≥ 140 mg/dL (7.8 mmol/L) after 100 g oral glucose tolerance test
- ADIPS (1998)—fasting plasma glucose ≥ 5.5 mmol/L or two hour glucose ≥ 8.0 mmol/L after 75 g oral glucose tolerance test
- Carpenter and Coustan—two or more raised values during oral glucose tolerance test: fasting plasma glucose ≥ 95 mg/dL (5.3 mmol/L), one hour glucose ≥ 180 mg/dL (10.0 mmol/L), two hour glucose ≥ 155 mg/dL (8.6 mmol/L), three hour glucose ≥ 140 mg/dL (7.8 mmol/L) after 100 g oral glucose tolerance test
- Contemporary criteria—one or more raised values during oral glucose tolerance test: fasting plasma glucose > 4.8 mmol/L, one hour glucose > 10.0 mmol/L, two hour glucose > 8.7 mmol/L until September 2001; since September 2001 fasting plasma glucose > 4.8 mmol/L, one hour glucose > 11.2 mmol/L, two hour glucose > 9.9 mmol/L after 75 g oral glucose tolerance test
- EASD—two hour glucose ≥ 9 mmol/L after 75 g oral glucose tolerance test
- IADPSG—fasting plasma glucose ≥ 92 mg/dL (5.1 mmol/L), one hour glucose ≥ 180 mg/dL (10.0 mmol/L), two hour glucose ≥ 153 mg/dL (8.5 mmol/L) after 75 g oral glucose tolerance test
- NICE—patients were classified as having GDM or T2DM based on clinical codes used in the *Health Improvement Network* database that were based on the classification made by NICE 2008 and 2015 guidelines for diabetes in pregnancy:
 - NICE 2008—fasting plasma glucose ≥ 7.0 mmol/L or two hour glucose ≥ 7.8 mmol/L after 75 g oral glucose tolerance test
 - NICE 2015: fasting plasma glucose ≥ 5.6 mmol/L or two hour glucose ≥ 7.8 mmol/L after 75 g oral glucose tolerance test
- NDDG (1979)—fasting plasma glucose ≥ 105 mg/dL (5.8 mmol/L), one hour glucose ≥ 190 mg/dL (10.6 mmol/L), two hour glucose ≥ 165 mg/dL (9.2 mmol/L), three hour glucose ≥ 145 mg/dL (8.0 mmol/L) after 100 g oral glucose tolerance test
- Second and third workshop conference on GDM—two or more raised values during oral glucose tolerance test: fasting plasma glucose ≥ 5.8 mmol/L, one hour glucose ≥ 10.6 mmol/L, two hour glucose ≥ 9.2 mmol/L, three hour glucose ≥ 8.1 mmol/L after 100 g oral glucose tolerance test
- WHO (1985)—fasting plasma glucose ≥ 7.0 mmol/L or two hour glucose ≥ 11.1 mmol/L after 75 g oral glucose tolerance test
- WHO (1999)—fasting plasma glucose ≥ 7.0 mmol/L, two hour glucose ≥ 7.8 mmol/L after 75 g oral glucose tolerance test

Criteria for type 2 diabetes mellitus (T2DM)

- ADA (1997)—fasting plasma glucose ≥ 7.0 mmol/L or two hour glucose ≥ 11.1 mmol/L after 75 g oral glucose tolerance test
- CDA (2008)—fasting plasma glucose ≥ 7.0 mmol/L or two hour glucose ≥ 11.1 mmol/L after 75 g oral glucose tolerance test
- CDA (2013)—fasting plasma glucose ≥ 7.0 mmol/L or two hour glucose ≥ 11.1 mmol/L or HbA1c $\geq 6.5\%$ (48 mmol/mol)
- Local criteria—HbA1c $\geq 7.25\%$ (55.7 mmol/mol) or glucose ≥ 11.1 mmol/L
- NICE: patients were classified as having GDM or T2DM based on clinical codes used in the *Health Improvement Network* database that were based on the classification made by the NICE 2008 and 2015 guidelines for diabetes in pregnancy:
 - NICE 2008—fasting plasma glucose ≥ 7.0 mmol/L, two hour glucose ≥ 11.1 mmol/L, diagnosis with the 75 g oral glucose tolerance test
 - NICE 2015—fasting plasma glucose ≥ 7.0 mmol/L (receive further testing to confirm T2DM), HbA1c $\geq 6.5\%$ (48.0 mmol/mol)
- WHO (1985)—fasting plasma glucose ≥ 7.8 mmol/L or two hour glucose ≥ 11.1 mmol/L after 75 g oral glucose tolerance test
- WHO (1998)—fasting plasma glucose ≥ 7.0 mmol/L, two hour glucose ≥ 11.1 mmol/L after 75 g oral glucose tolerance test

No indication of publication bias was found, with both Begg's and Egger's tests being statistically non-significant ($P=0.58$ and $P=0.90$, respectively). The funnel plot of the 20 studies included in the meta-analysis is provided in supplementary figure S4.

Discussion**Principal findings**

The results of this systematic review and meta-analysis suggest that women with a history of GDM are almost 10 times more likely to develop T2DM than those with a normoglycaemic pregnancy. The magnitude of this risk is consistent with evidence that the two conditions share common pathogenic mechanisms and risk factors,⁵² suggesting that GDM could potentially serve as a predictor for future development of T2DM.

Strengths and limitations

Among the strengths of this study are the substantial number of recently published studies, assessing a large total number of individuals, and with long term follow-up, ranging from one to 25 years. Guidelines for the screening, diagnosis, and follow-up care for women

with GDM have changed in a number of countries over the past decade, while the overall prevalence of T2DM has increased. This review provides up-to-date results in contemporary populations.

Nevertheless, this study had several limitations. Owing to limited resources, we had to exclude studies not published in English. Furthermore, owing to a lack of information on family history of diabetes or parity, we were not able to examine the potential effect of these factors on the incidence of T2DM. Additionally, while several primary studies adequately reported the ethnicity of study participants, others provided little or no information and had to be grouped using broad ethnic categories. Thus, we could not investigate progression in ethnic subgroups, which could have been a cause of heterogeneity between studies. The subsequent risk of T2DM in both women with previous GDM and healthy controls was estimated using relative risks, as person years of follow-up were not reported for every study. Thus we were not able to calculate incidence rate ratios consistently across studies. A further limitation was that although we assessed the cumulative incidence by study length of follow-up, it

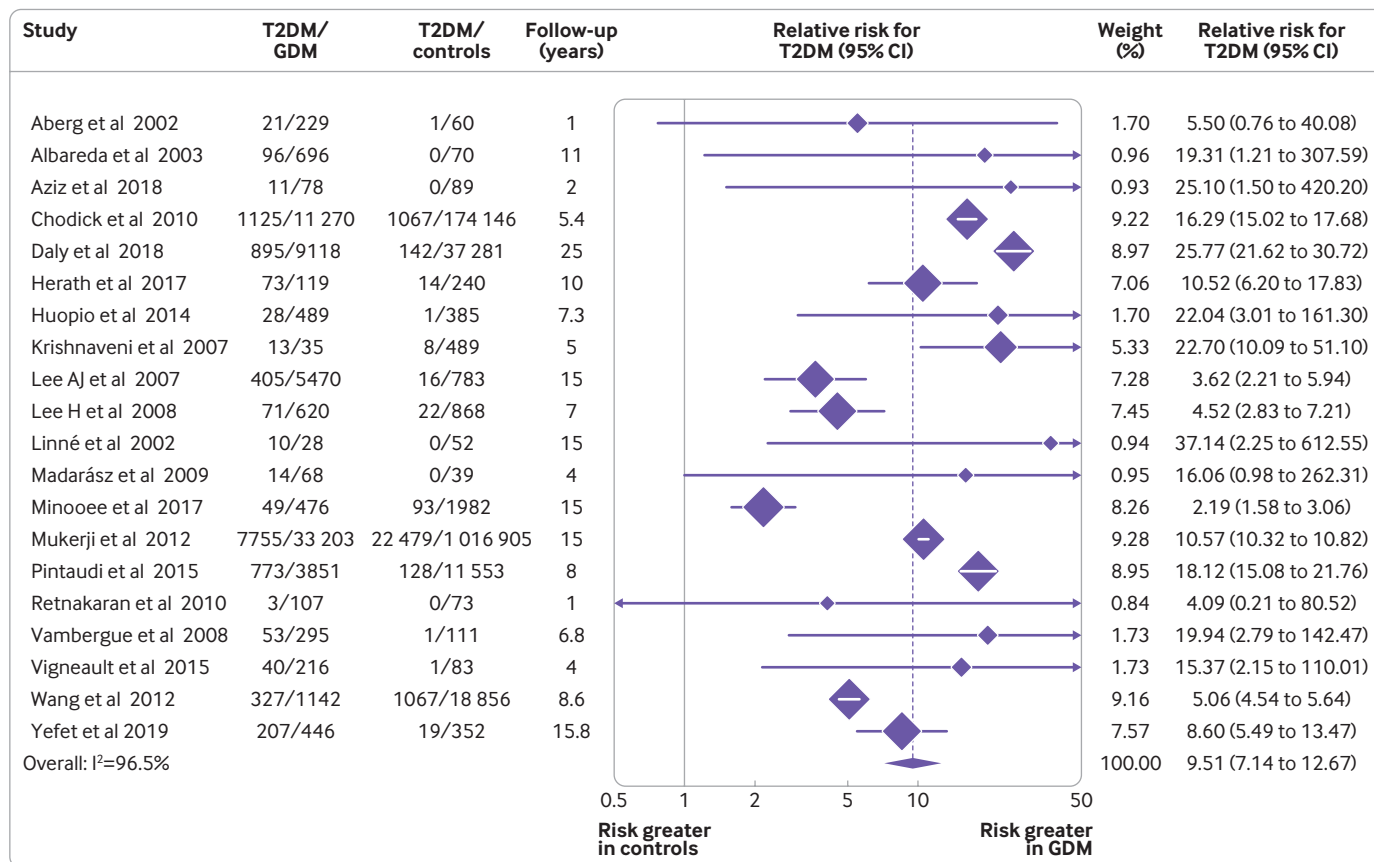


Fig 2 | Relative risk of type 2 diabetes mellitus (T2DM) in women with gestational diabetes mellitus (GDM) compared with healthy controls

was difficult to reach conclusions about the timing of T2DM onset using study level data, as the cumulative incidence was known only at the end of the study and not when the events occurred. A more accurate assessment of cumulative incidence would require individual patient data, in a cohort where regular screening took place. In addition, the prespecified meta-regression analyses carried out to assess heterogeneity between studies lack power to identify associations between variables, as they are limited to the use of study level data. Overall, we were unable to identify the main sources of heterogeneity in our effect estimates, and a more in-depth analysis could have been performed if individual patient level data had been available.

To assess the quality of our included studies, we used the Newcastle-Ottawa scale. The validity of this scale has been questioned by previous studies, because it gives little attention to study confounders.²² Although the Risk Of Bias In Non-randomised Studies - of Interventions (ROBINS-I) has been suggested as a high quality, validated scale,⁵³ we were not able to use it for this meta-analysis, because our primary studies did not involve any interventions. To deal with this problem, we carefully evaluated potential confounders in primary studies, and listed them thoroughly. Most of the studies included were at low risk of bias, with only one study⁴³ potentially at high risk of bias. The funnel plot and statistical tests

performed showed no evidence of publication bias in our meta-analysis.

Comparison with other studies

Our findings are similar to results from previously published systematic reviews, suggesting a higher incidence of T2DM in women with GDM than in those unaffected, with the overall 95% confidence interval obtained by Bellamy et al overlapping with the one we identified.^{12 13} In this systematic review and meta-analysis, we incorporated evidence from the most recent studies and identified a higher estimated risk of T2DM than found previously.

Kim et al and Bellamy et al assessed studies published between 1965 and 2001 and between 1960 and 2009, respectively.^{12 13} As screening criteria have changed, with the cut-off level of fasting plasma glucose for T2DM being lowered from 7.8 to 7.0 mmol/L in 1997,⁵⁴ this could have led to an underestimation of the overall risk of subsequent T2DM in women with previous GDM. Some of our included studies, however, involved both women diagnosed with the criteria used before 1997, and women diagnosed with the revised criteria.^{31 34-36 38 41 43 48 49} Additionally, previous systematic reviews have focused on a specific ethnic group,¹⁴ or included studies with short term follow-up (≤ 6 months).^{12 14} Kim et al also showed a higher cumulative incidence of T2DM in the first five

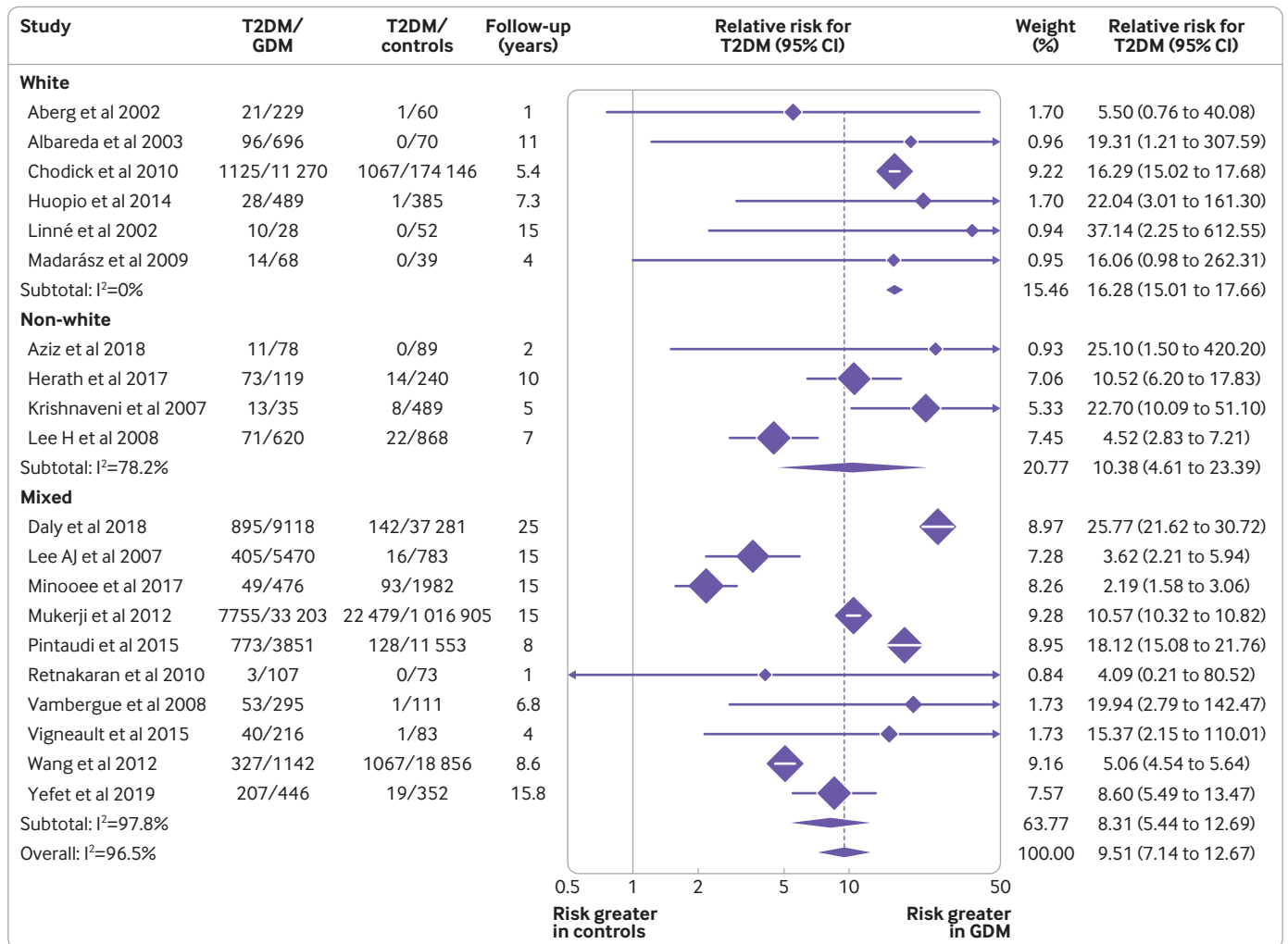


Fig 3 | Relative risk of type 2 diabetes mellitus (T2DM) in women with gestational diabetes mellitus (GDM) and controls by study level ethnicity

postpartum years, which then seemed to plateau in further years.¹² In our systematic review, the cumulative incidence of T2DM was seen to increase steadily over time, with the highest incidence being seen in those studies with the longest follow-up.

This analysis did not show a statistically significant difference in the relative risk of T2DM in different ethnic groups. This result could be due to a lack of power to detect any potential differences, consistent with previous research showing that after adjusting for covariates, progression to T2DM between different ethnic groups seemed similar in all women with previous GDM.¹² Similarly, Bellamy et al found no difference in the overall relative risk between the different ethnic groups of the included studies.¹³

Implications for research and clinical practice

The substantially higher risk of subsequent progression to T2DM in women diagnosed with GDM identified in this systematic review is perhaps not surprising, considering the poor postpartum screening uptake in this population, and lack of structured postpartum preventive care.⁵⁵ Some of the barriers highlighted by previous research include

poor communication between clinicians and patients, and between healthcare professionals in primary and secondary care, lack of awareness of T2DM risk due to poor patient education, and time restrictions due to maternal duties.⁵⁶ Our results could increase awareness in patients of the need to attend postpartum screening, and in healthcare professionals of the need to use patient centred strategies to improve screening uptake. Future research could further investigate the timing of onset of T2DM in these populations, using individual patient data, to provide a more accurate estimate of time.

The long term health benefits associated with the adoption of lifestyle and pharmacological interventions aimed at preventing the onset of T2DM in women with GDM have been well recognised.⁵⁷ Nevertheless, more up-to-date large randomised controlled trials are needed to investigate the effectiveness of those interventions in ethnically diverse populations, and with longer follow-up, to support the generalisability and external validity of these results. Finally, the cost effectiveness of these interventions should be considered to promote their adoption by different healthcare systems across the world.

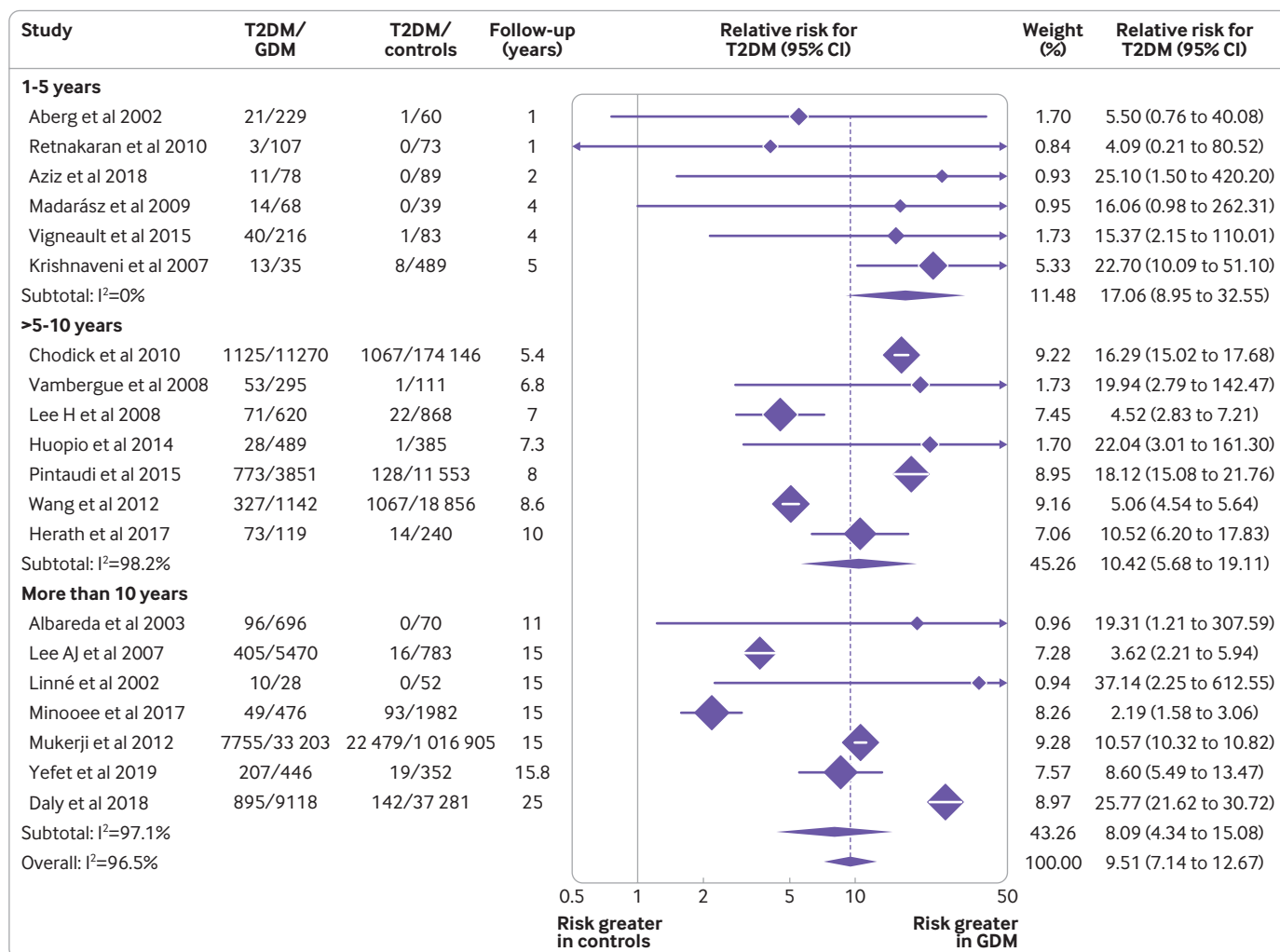


Fig 4 | Relative risk of type 2 diabetes mellitus (T2DM) in women with gestational diabetes mellitus (GDM) and controls based on study length of follow-up

	No of contributing studies	Study length (years; mean (range))	GDM (%; 95% CI)	Controls (%; 95% CI)
Ethnicity:				
White	6	7.28 (1.00-15.00)	9.91 (9.39 to 10.42)	0.61 (0.58 to 0.65)
Mixed	10	11.44 (1.00-25.00)	16.46 (16.16 to 16.77)	1.91 (1.88 to 1.93)
Non-white	4	6.00 (2.00-10.00)	15.58 (13.30 to 17.86)	2.02 (1.35 to 2.69)
Study follow-up length (years):				
1-5	6	2.83 (1.00-5.00)	9.22 (7.19 to 11.26)	1.19 (0.46 to 1.93)
>5-10	7	7.61 (5.40-10.00)	12.06 (11.59 to 12.53)	0.69 (0.66 to 0.73)
>10	7	15.97 (11.00-25.00)	16.15 (15.83 to 16.47)	1.90 (1.87 to 1.92)

Conclusions

In summary, women with a history of GDM have a nearly 10-fold increased risk of developing T2DM, in comparison with those with a normoglycaemic pregnancy. These high rates of progression suggest an urgent need to promote postpartum screening among

these women and to encourage them to adopt dietary, lifestyle, and pharmacological interventions to prevent or delay the onset of T2DM. Future studies should examine strategies to improve screening uptake and evaluate the effectiveness and cost effectiveness of preventive interventions, across heterogeneous populations and over long periods.

Study level variables	Coefficient (95% CI)	P value	I ² (%)
Age (mean)	1.02 (0.93 to 1.12)	0.67	96.00
Body mass index (mean)	0.97 (0.87 to 1.08)	0.52	80.43
Publication year	1.01 (0.92 to 1.10)	0.87	96.66
Length of follow-up (mean)	1.00 (0.93 to 1.07)	0.94	96.67

Contributors: EV, KK, and CLG conceived the idea of the study; EV and SCA screened the studies and undertook data extraction; EV carried out the statistical analysis; CLG supervised the analysis. EV, CLG, KK, BKT, and MJD interpreted the findings; and EV drafted the manuscript. CLG, KK, MJD, and BKT critically reviewed the manuscript and EV revised the manuscript for final submission. All authors have approved the final draft of the manuscript. EV is guarantor. EV accepts full

responsibility for the work and the conduct of the study, had access to the data, and controlled the decision to publish. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This report is the independent research of EV, supported by the National Institute for Health Research (NIHR) Applied Research Collaboration-East Midlands as part of a PhD project. This research is also supported by the NIHR Leicester Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NIHR, National Health Service, or the Department of Health and Social Care. The funders had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: support from the NIHR Applied Research Collaboration East Midlands and NIHR Leicester Biomedical Research Centre for the submitted work. KK has acted as a consultant and speaker for Novartis, Novo Nordisk, Sanofi-Aventis, Lilly, Servier, and Merck Sharp and Dohme; has received grants in support of investigator and investigator initiated trials from Novartis, Novo Nordisk, Sanofi-Aventis, Lilly, Pfizer, Boehringer Ingelheim, and Merck Sharp and Dohme; and has received funds for research, honoraria for speaking at meetings and has served on advisory boards for Lilly, Sanofi-Aventis, Merck Sharp and Dohme, and Novo Nordisk. MJD has acted as consultant, advisory board member and speaker for Novo Nordisk, Sanofi-Aventis, Lilly, Merck Sharp and Dohme, Boehringer Ingelheim, AstraZeneca, and Janssen; as a speaker for Mitsubishi Tanabe Pharma Corporation; and has received grants in support of investigator and investigator initiated trials from Novo Nordisk, Sanofi-Aventis, and Lilly. All other authors declare no competing interests, or activities that could appear to have influenced the submitted work.

Ethical approval: Ethical approval was not required.

Data sharing: Data are available upon reasonable request. Technical appendix, statistical code, and dataset available from the corresponding author at ev63@le.ac.uk.

The lead author (EV) affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Dissemination to participants and related patient and public communities: The results of this study will be disseminated to the Royal College of General Practitioners and the diabetes prevention programme committee, of which KK is a member. The authors will additionally issue a press release in collaboration with the BMJ. These findings will inform the Baby Steps study run by the Diabetes Research Centre, which aims to implement a structured group education programme for women with a history of gestational diabetes. Finally, results will also be disseminated to national and international diabetes conferences. Where possible, results of this systematic review and meta-analysis will be disseminated to the patient community and to regional and national committees through the study authors.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

- American Diabetes Association. Classification and diagnosis of diabetes: standards of medical care in diabetes 2019. *Diabetes Care* 2019;42(Suppl 1):S13-28. doi:10.2337/dc19-S002.
- International Diabetes Federation. *IDF Diabetes Atlas - 2019*. International Diabetes Federation, 2019.
- Goueslard K, Cottenet J, Mariet AS, Sagot P, Petit JM, Quantin C. Early screening for type 2 diabetes following gestational diabetes mellitus in France: hardly any impact of the 2010 guidelines. *Acta Diabetol* 2017;54:645-51. doi:10.1007/s00592-017-0986-x.
- Kim C, Tabaei BP, Burke R, et al. Missed opportunities for type 2 diabetes mellitus screening among women with a history of gestational diabetes mellitus. *Am J Public Health* 2006;96:1643-8. doi:10.2105/AJPH.2005.065722.
- Kwong S, Mitchell RS, Senior PA, Chik CL. Postpartum diabetes screening: adherence rate and the performance of fasting plasma glucose versus oral glucose tolerance test. *Diabetes Care* 2009;32:2242-4. doi:10.2337/dc09-0900.
- Blatt AJ, Nakamoto JM, Kaufman HW. Gaps in diabetes screening during pregnancy and postpartum. *Obstet Gynecol* 2011;117:61-8. doi:10.1097/AOG.0b013e3181fe424b.
- Kim C, McEwen LN, Piette JD, Goewey J, Ferrara A, Walker EA. Risk perception for diabetes among women with histories of gestational diabetes mellitus. *Diabetes Care* 2007;30:2281-6. doi:10.2337/dc07-0618.
- Sterne VL, Logan T, Palmer MA. Factors affecting attendance at postpartum diabetes screening in women with gestational diabetes mellitus. *Pract Diabetes Int* 2011;28:64-9. doi:10.1002/pdi.1559.
- Minsart AF, Vander Maelen A, Fontaine V, Kirkpatrick C. Social, medical and self-perceived factors influencing postpartum screening of diabetes after gestational diabetes. *J Obstet Gynaecol* 2014;34:8-12. doi:10.3109/01443615.2013.826639.
- Seidu S, Khunti K. Non-adherence to diabetes guidelines in primary care - the enemy of evidence-based practice. *Diabetes Res Clin Pract* 2012;95:301-2. doi:10.1016/j.diabres.2012.01.015.
- Almario CV, Ecker T, Moroz LA, Bucovetsky L, Berghella V, Baxter JK. Obstetricians seldom provide postpartum diabetes screening for women with gestational diabetes. *Am J Obstet Gynecol* 2008;198:528.e1-5. doi:10.1016/j.ajog.2007.11.001.
- Kim C, Newton KM, Knopp RH. Gestational diabetes and the incidence of type 2 diabetes: a systematic review. *Diabetes Care* 2002;25:1862-8. doi:10.2337/diacare.25.10.1862.
- Bellamy L, Casas JP, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. *Lancet* 2009;373:1773-9. doi:10.1016/S0140-6736(09)60731-5.
- Nouhjah S, Shahbazian H, Amoori N, et al. Postpartum screening practices, progression to abnormal glucose tolerance and its related risk factors in Asian women with a known history of gestational diabetes: a systematic review and meta-analysis. *Diabetes Metab Syndr* 2017;11(Suppl 2):S703-12. doi:10.1016/j.dsx.2017.05.002.
- Ferrara A. Increasing prevalence of gestational diabetes mellitus: a public health perspective. *Diabetes Care* 2007;30(Suppl 2):S141-6. doi:10.2337/dc07-s206.
- Li-Zhen L, Yun X, Xiao-Dong Z, et al. Evaluation of guidelines on the screening and diagnosis of gestational diabetes mellitus: systematic review. *BMJ Open* 2019;9:e023014. doi:10.1136/bmjopen-2018-023014.
- World Health Organization. *Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus*. World Health Organization, 2011.
- Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;62:e1-34. doi:10.1016/j.jclinepi.2009.06.006.
- Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;283:2008-12. doi:10.1001/jama.283.15.2008.
- Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions*. Cochrane Collaboration, 2011.
- World Health Organization. *Definition, diagnosis and classification of diabetes mellitus and its complications: report of a WHO consultation. Part 1, Diagnosis and classification of diabetes mellitus*. World Health Organization, 1999.
- Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol* 2010;25:603-5. doi:10.1007/s10654-010-9491-z.
- DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-88. doi:10.1016/0197-2456(86)90046-2.
- Altman DG. *Practical statistics for medical research*. Chapman and Hall, 1997.
- Tehrani FR, Hashemi S, Hashemina M, Azizi F. Follow-up of women with gestational diabetes in the Tehran Lipid and Glucose Study (TLGS): a population-based cohort study. *J Obstet Gynaecol Res* 2012;38:698-704. doi:10.1111/j.1447-0756.2011.01767.x.
- Hakkarainen H, Huopio H, Cederberg H, Pääkkönen M, Voutilainen R, Heinonen S. Post-challenge glycemia during pregnancy as a marker of future risk of type 2 diabetes: a prospective cohort study. *Gynecol Endocrinol* 2015;31:573-7. doi:10.3109/09513590.2015.1032926.
- Minooee S, Ramezani Tehrani F, Rahmati M, Mansournia MA, Azizi F. Diabetes incidence and influencing factors in women with and without gestational diabetes mellitus: a 15 year population-based follow-up cohort study. *Diabetes Res Clin Pract* 2017;128:24-31. doi:10.1016/j.diabres.2017.04.003.
- Huopio H, Hakkarainen H, Pääkkönen M, et al. Long-term changes in glucose metabolism after gestational diabetes: a double cohort study. *BMC Pregnancy Childbirth* 2014;14:296. doi:10.1186/1471-2393-14-296.
- Retnakaran R, Luo J, Shah BR. Gestational diabetes in young women predicts future risk of serious liver disease. *Diabetologia* 2019;62:306-10. doi:10.1007/s00125-018-4775-z.
- Feig DS, Zinman B, Wang X, Hux JE. Risk of development of diabetes mellitus after diagnosis of gestational diabetes. *CMAJ* 2008;179:229-34. doi:10.1503/cmaj.080012.

- 31 Mukerji G, Chiu M, Shah BR. Impact of gestational diabetes on the risk of diabetes following pregnancy among Chinese and South Asian women. *Diabetologia* 2012;55:2148-53. doi:10.1007/s00125-012-2549-6.
- 32 Kousta E, Efstathiadou Z, Lawrence NJ, et al. The impact of ethnicity on glucose regulation and the metabolic syndrome following gestational diabetes. *Diabetologia* 2006;49:36-40. doi:10.1007/s00125-005-0058-6.
- 33 Vambergue A, Dognin C, Boulogne A, Réjou MC, Biaisque S, Fontaine P. Increasing incidence of abnormal glucose tolerance in women with prior abnormal glucose tolerance during pregnancy: DIAGEST 2 study. *Diabet Med* 2008;25:58-64. doi:10.1111/j.1464-5491.2007.02306.x.
- 34 Chodick G, Elchalal U, Sella T, et al. The risk of overt diabetes mellitus among women with gestational diabetes: a population-based study. *Diabet Med* 2010;27:779-85. doi:10.1111/j.1464-5491.2010.02995.x.
- 35 Albareda M, Caballero A, Badell G, et al. Diabetes and abnormal glucose tolerance in women with previous gestational diabetes. *Diabetes Care* 2003;26:1199-205. doi:10.2337/diacare.26.4.1199.
- 36 Aberg AEB, Jönsson EK, Eskilsson I, Landin-Olsson M, Frid AH. Predictive factors of developing diabetes mellitus in women with gestational diabetes. *Acta Obstet Gynecol Scand* 2002;81:11-6. doi:10.1046/j.0001-6349.2001.00000.x.
- 37 Aziz S, Munim TF, Fatima SS. Post-partum follow-up of women with gestational diabetes mellitus: effectiveness, determinants, and barriers. *J Matern Fetal Neonatal Med* 2018;31:1607-12. doi:10.1080/014767058.2017.1321630.
- 38 Daly B, Toulis KA, Thomas N, et al. Increased risk of ischemic heart disease, hypertension, and type 2 diabetes in women with previous gestational diabetes mellitus, a target group in general practice for preventive interventions: a population-based cohort study [*PLoS Med* 2019;16:e1002881]. *PLoS Med* 2018;15:e1002488. doi:10.1371/journal.pmed.1002488.
- 39 Herath H, Herath R, Wickremasinghe R. Gestational diabetes mellitus and risk of type 2 diabetes 10 years after the index pregnancy in Sri Lankan women—a community based retrospective cohort study. *PLoS One* 2017;12:e0179647. doi:10.1371/journal.pone.0179647.
- 40 Krishnaveni GV, Hill JC, Veena SR, et al. Gestational diabetes and the incidence of diabetes in the 5 years following the index pregnancy in South Indian women. *Diabetes Res Clin Pract* 2007;78:398-404. doi:10.1016/j.diabres.2007.06.002.
- 41 Lee AJ, Hiscok RJ, Wein P, Walker SP, Permezel M. Gestational diabetes mellitus: clinical predictors and long-term risk of developing type 2 diabetes: a retrospective cohort study using survival analysis. *Diabetes Care* 2007;30:878-83. doi:10.2337/dc06-1816.
- 42 Lee H, Jang HC, Park HK, Metzger BE, Cho NH. Prevalence of type 2 diabetes among women with a previous history of gestational diabetes mellitus. *Diabetes Res Clin Pract* 2008;81:124-9. doi:10.1016/j.diabres.2008.02.017.
- 43 Linné Y, Barkeling B, Rössner S. Natural course of gestational diabetes mellitus: long term follow up of women in the SPAWN study. *BJOG* 2002;109:1227-31. doi:10.1046/j.1471-0528.2002.01373.x.
- 44 Madarász E, Tamás G, Tabák ÁG, Kerényi Z. Carbohydrate metabolism and cardiovascular risk factors 4 years after a pregnancy complicated by gestational diabetes. *Diabetes Res Clin Pract* 2009;85:197-202. doi:10.1016/j.diabres.2009.05.001.
- 45 Pintauro B, Lucisano G, Pellegrini F, et al. The long-term effects of stillbirth on women with and without gestational diabetes: a population-based cohort study. *Diabetologia* 2015;58:67-74. doi:10.1007/s00125-014-3403-9.
- 46 Retnakaran R, Qi Y, Sermer M, Connelly PW, Hanley AJ, Zimman B. Beta-cell function declines within the first year postpartum in women with recent glucose intolerance in pregnancy. *Diabetes Care* 2010;33:1798-804. doi:10.2337/dc10-0351.
- 47 Vigneault J, Lemieux S, Garneau V, Weisnagel SJ, Tcherno A, Robitaille J. Association between metabolic deteriorations and prior gestational diabetes according to weight status. *Obesity (Silver Spring)* 2015;23:345-50. doi:10.1002/oby.20940.
- 48 Wang Y, Chen L, Horswell R, et al. Racial differences in the association between gestational diabetes mellitus and risk of type 2 diabetes. *J Womens Health (Larchmt)* 2012;21:628-33. doi:10.1089/jwh.2011.3318.
- 49 Yefet E, Schwartz N, Sliman B, Ishay A, Nachum Z. Good glycemic control of gestational diabetes mellitus is associated with the attenuation of future maternal cardiovascular risk: a retrospective cohort study. *Cardiovasc Diabetol* 2019;18:75. doi:10.1186/s12933-019-0881-6.
- 50 Wang B, An X, Shi X, Zhang JA. Management of endocrine disease: suicide risk in patients with diabetes: a systematic review and meta-analysis. *Eur J Endocrinol* 2017;177:R169-81. doi:10.1530/EJE-16-0952.
- 51 Bonovas S, Fiorino G, Lytras T, Nikolopoulos G, Peyrin-Birolet L, Danese S. Systematic review with meta-analysis: use of 5-aminosalicylates and risk of colorectal neoplasia in patients with inflammatory bowel disease. *Aliment Pharmacol Ther* 2017;45:1179-92. doi:10.1111/apt.14023.
- 52 Buchanan TA, Xiang A, Kjos SL, Watanabe R. What is gestational diabetes? *Diabetes Care* 2007;30(Suppl 2):S105-11. doi:10.2337/dc07-s201.
- 53 Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;355:i4919. doi:10.1136/bmj.i4919.
- 54 Kousta E, Lawrence NJ, Penny A, et al. Implications of new diagnostic criteria for abnormal glucose homeostasis in women with previous gestational diabetes. *Diabetes Care* 1999;22:933-7. doi:10.2337/diacare.22.6.933.
- 55 McGovern A, Butler L, Jones S, et al. Diabetes screening after gestational diabetes in England: a quantitative retrospective cohort study. *Br J Gen Pract* 2014;64:e17-23. doi:10.3399/bjgp14X676410.
- 56 Sanderson H, Loveman E, Colquitt J, Royle P, Waugh N, Tan BK. Improving uptake of postnatal checking of blood glucose in women who had gestational diabetes mellitus in universal healthcare settings: a systematic review. *J Clin Med* 2018;8:4. doi:10.3390/jcm8010004.
- 57 Morton S, Kirkwood S, Thangaratinam S. Interventions to modify the progression to type 2 diabetes mellitus in women with gestational diabetes: a systematic review of literature. *Curr Opin Obstet Gynecol* 2014;26:476-86. doi:10.1097/GCO.0000000000000127.

Web appendix: Supplementary material