Original Article
Metabolic Risk/Epidemiology

Diabetes Metab J Published online Mar 8, 2022
https://doi.org/10.4093/dmj.2021.0178
pISSN 2233-6079 · eISSN 2233-6087

dmj
DIABETES & METABOLISM JOURNAL

Effect of Different Types of Diagnostic Criteria for Gestational Diabetes Mellitus on Adverse Neonatal Outcomes: A Systematic Review, Meta-Analysis, and Meta-Regression

Fahimeh Ramezani Tehrani[1], Marzieh Saei Ghare Naz[1], Razieh Bidhendi-Yarandi[2], Samira Behboudi-Gandevani[3]

[1]Reproductive Endocrinology Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran,
[2]Department of Biostatistics, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran,
[3]Faculty of Nursing and Health Sciences, Nord University, Bodo, Norway

Background: Evidence supporting various diagnostic criteria for diagnose gestational diabetes mellitus (GDM) are consensus-based, needs for additional evidence related to outcomes. Therefore, the aim of this systematic-review and meta-analysis was to assess the impact of different GDM diagnostic-criteria on the risk of adverse-neonatal-outcomes.
Methods: Electronic databases including Scopus, PubMed, and Web of Sciences were searched to retrieve English original, population-based studies with the universal GDM screening approach, up to January-2020. GDM diagnostic criteria were classified in seven groups and International Association of the Diabetes and Pregnancy Study Groups (IADPSG) was considered as reference one. We used the Mantel–Haenszel method to calculate the pooled odds of events. The possibility of publication bias was examined by Begg's test.
Results: A total of 55 population-based studies consisting of 1,604,391 pregnant women with GDM and 7,770,855 non-GDM counterparts were included. Results showed that in all diagnostic-criteria subgroups, the risk of adverse neonatal outcomes including macrosomia, hyperbilirubinemia, respiratory distress syndrome, neonatal hypoglycemia, neonatal intensive care unit admission, preterm birth, and birth-trauma were significantly higher than the non-GDM counterparts were significantly higher than non-GDM counterparts. Meta-regression analysis revealed that the magnitude of neonatal risks in all diagnostic-criteria subgroups are similar.
Conclusion: Our results showed that the risk of adverse-neonatal-outcome increased among women with GDM, but the magnitude of risk was not different among those women who were diagnosed through more or less intensive strategies. These findings may help health-care-providers and policy-makers to select the most cost-effective approach for the screening of GDM among pregnant women.

Keywords: Diabetes, gestational; Meta-analysis; Pregnancy complications

INTRODUCTION

Gestational diabetes mellitus (GDM) is a globally rising health problem [1]. According to American Diabetes Association (ADA), GDM is defined as "diabetes diagnosed in the second or third trimester of pregnancy that was not clearly overt diabetes prior to gestation" [2]. The prevalence of GDM among various population is varied between 4% and 15% [3].

GDM results from impaired secretory response of pancreatic β-cell to increased maternal insulin demands during pregnan-

Corresponding author: Samira Behboudi-Gandevani
https://orcid.org/0000-0003-3526-640X
Faculty of Nursing and Health Sciences, Nord University, 8049 Bodø, Norway
E-mail: samira.behboudi-gandevani@nord.no

Received: Jul. 31, 2021; Accepted: Oct. 2, 2021

cy [4]. Variety of factors including family history of diabetes, previous history of macrocosmic babies, higher first trimester body mass index and older maternal age lead to increased risk of developing GDM [5,6]. It is well acknowledge that GDM is associated with an increased short and long-term risk of complications for mothers and their babies [7-10].

Despite health consensus recommend various diagnostic criteria for GDM, there is no consensus about the optimal screening and diagnosis criteria [11]. In 2010 (a decade ago), the International Association of the Diabetes and Pregnancy Study Group (IADPSG) [12] provided stringent threshold for GDM diagnosis by one-step 75 g oral glucose tolerance tests based on the results of the observational Hyperglycemia and Adverse Pregnancy Outcomes (HAPO) study [13], later on, some other expert professional organizations including International Federation of Gynecology and Obstetrics (FIGO) [14], ADA, and World Health Organization (WHO) [15] supported the recommendation of IADPSG. However, the American College of Obstetricians and Gynecologists (ACOG) has always endorsed the two-step approach to GDM [16]. Furthermore some countries follow the national guideline with different diagnostic approach and glucose thresholds [17-24]. The main point is that the evidence supporting these endorsements are consensus-based, and both main organizations of IADPSG and ACOG note the need for additional evidence related to outcomes [2,25].

In addition, although there is a clear linear relationship between maternal hyperglycemia and maternal and perinatal outcomes, the effects of identifying and treating milder cases of gestational diabetes on these outcomes are not known yet [25-27]. By conducting this meta-analysis, we tried to fill the gap of knowledge, based on available evidence to find the impact of different gestational-diabetes (GDM) diagnostic-criteria on the risk of adverse-neonatal-outcomes.

METHODS

This study approved by ethics committee of the Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran (IR.SBMU.ENDOCRINE.REC.1399.076). Informed consent was waived by the board.

The present review study was reported based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [28] to examine the following objectives:

(1) To examine the pooled odds of adverse neonatal outcomes among participants with GDM compared to non-GDM groups

(2) To examine the pooled odds of adverse neonatal outcomes among participants with GDM compared to non-GDM counterparts, based on the various screening and diagnostic criteria for GDM

Inclusion criteria

Studies were entered into the analysis if they were: (1) universally screened the GDM; (2) provided accurate screening and diagnosis criteria; (3) being population-based design; (4) provided the one of the short-term neonatal outcomes of in both pregnant women with GDM and non-GDM. Studies were excluded if they were: reviews, letter to editor, meeting abstracts, case reports. There are no restrictions regarding country, age, race and other demographic characteristics of counterparts.

Search strategy

A comprehensive systematic search up to January 2020 was performed in the electronic databases including PubMed, Web of Sciences, and Scopus to retrieve relevant English publications based on the combination the keywords as follows: adverse pregnancy outcome, pregnancy outcome, pregnancy complication, small for gestational age (SGA), macrosomia, large for gestational age (LGA), neonatal distress, respiratory distress syndrome (RDS), neonatal RDS, neonatal intensive care unit (NICU) admission, NICU, preterm, hyperbilirubinemia, stillbirth, neonatal hypoglycemia, birth trauma, shoulder dystocia, bone fracture, GDM, gestational diabetes, pregnancy-induced diabetes, glucose intolerances, and impaired glucose tolerances.

Study selection and data extraction

The title and abstract of records screened by two investigators (M.S.G. and S.B.G.) independently for determining final eligibility criteria. Any disagreements were discussed by two researchers and an another investigator (F.R.T.) until consensus was achieved. Two authors (M.S.G. and R.B.Y.) applied data extraction. Data were extracted from full text of studies including name of first author, country, years of publication, sample size, percent/number of events related to the each outcomes, diagnostic criteria for GDM, and population characteristics.

Study subgroups

All included studies were classified in seven subgroups based

Table 1. Screening and diagnosis of gestational diabetes mellitus classifications

Classification	Glucose load	No. of impaired value for diagnosis	Threshold
Group 1	OGTT with 75 g 2-hr	1	>92, 180, and 153 mg/dL for fasting, 1 and 2 hr
Group 2	OGTT with 75 g 2-hr	1	>100 and 144 mg/dL for fasting and 2 hr
Group 3	OGTT with 75 g 2-hr	1	>110 and 140 mg/dL for fasting and 2 hr
Group 4	OGTT with 75 g 2-hr	1	>180 mg/dL for 2 hr
Group 5	GCT with 50 g 1-hr,		>140 mg/dL
	followed by OGTT with 100 g 3-hr	2	>95, 180, 155, and 140 mg/dL for fasting, 1, 2, and 3 hr
Group 6	GCT with 50 g 1-hr,		>140 mg/dL
	followed by OGTT with 100 g 3-hr	2	>105, 155, 165, and 145 mg/dL for fasting, 1, 2, and 3 hr
Group 7	OGTT with 100 g 3-hr	1	>120, 175, 155, and 140 mg/dL for fasting, 1, 2, and 3 hr

OGTT, oral glucose tolerance test; GCT, glucose challenge test.

on the screening and diagnosis approaches and closest value of blood glucose thresholds (Table 1).

Outcome measures

The main outcomes in this meta-analysis were nine separate neonatal short outcomes of SGA, preterm birth, LGA, stillbirth, macrosomia, hyperbilirubinemia, RDS, neonatal hypoglycemia, NICU admission, and one composite outcome of neonatal birth trauma (including bone fracture, shoulder dystocia, birth injury, and Erb's palsy).

Quality assessment and risk of bias

The methodological quality assessment of included studies was performed by two investigators (M.S.G. and S.B.G.) independently using the Newcastle-Ottawa Scale [10]. This scale is categorized into three dimensions including selection, comparability and outcomes. The scoring system (range, 0 to 9) is used to provide final judgment regarding the quality of included studies. Scores above 6, 3–5, and below 3 were interpreted as high, moderate, and low quality, respectively. The Cochrane Collaboration's tool was used for assessing the risk of bias of included studies [29]. Risk of bias of cross-sectional studies was performed in five domains including: bias in selection of cases and controls, control of prognostic variable and development of outcome also in cohort studies the risk of bias evaluation was performed in seven domains including selection, assessment of exposure and outcome, presence of outcome of interest at start of study, control of prognostic variables, presence or absence of prognostic factors and adequacy regarding follow-up of cohorts. The authors classified their judgment on the quality of each study into high risk, unclear risk, or low risk of bias.

Data analysis

All data analyses were performed in the STATA version 13 (STATA Inc., College Station, TX, USA). We used the Mantel–Haenszel method to calculate the pooled odds ratios (ORs) of events. The heterogeneity between studies was assessed using I^2 and Cochrane's Q test. Heterogeneous and non-heterogeneous results were analyzed using the random-effects and fixed effect model respectively [30]. The possibility of publication bias in the present study was examined by Begg's test. The trim and fill method was used to deal with publication bias [31]. Meta-regression was performed to investigate any potential source of heterogeneity among GDM diagnosis criteria (IADPSG as a reference group). All results reported in significance level of 0.05 with 95% confidence intervals (CIs).

RESULTS

Search results, study characteristics, and quality assessment

Fig. 1 shows the flow diagram of studies retrieval and study selection. In this review, 55 studies provide information of adverse neonatal outcomes of 1604,391 participants with GDM and 7,770,855 non-GDM participants. Details of the characteristics of included studies are presented in Supplementary Table 1. All studies were classified as high quality (Supplementary Tables 2 and 3) [32-86]. A total of 53 studies were prospective or retrospective cohorts [32-64,66-82,84-86] and two were

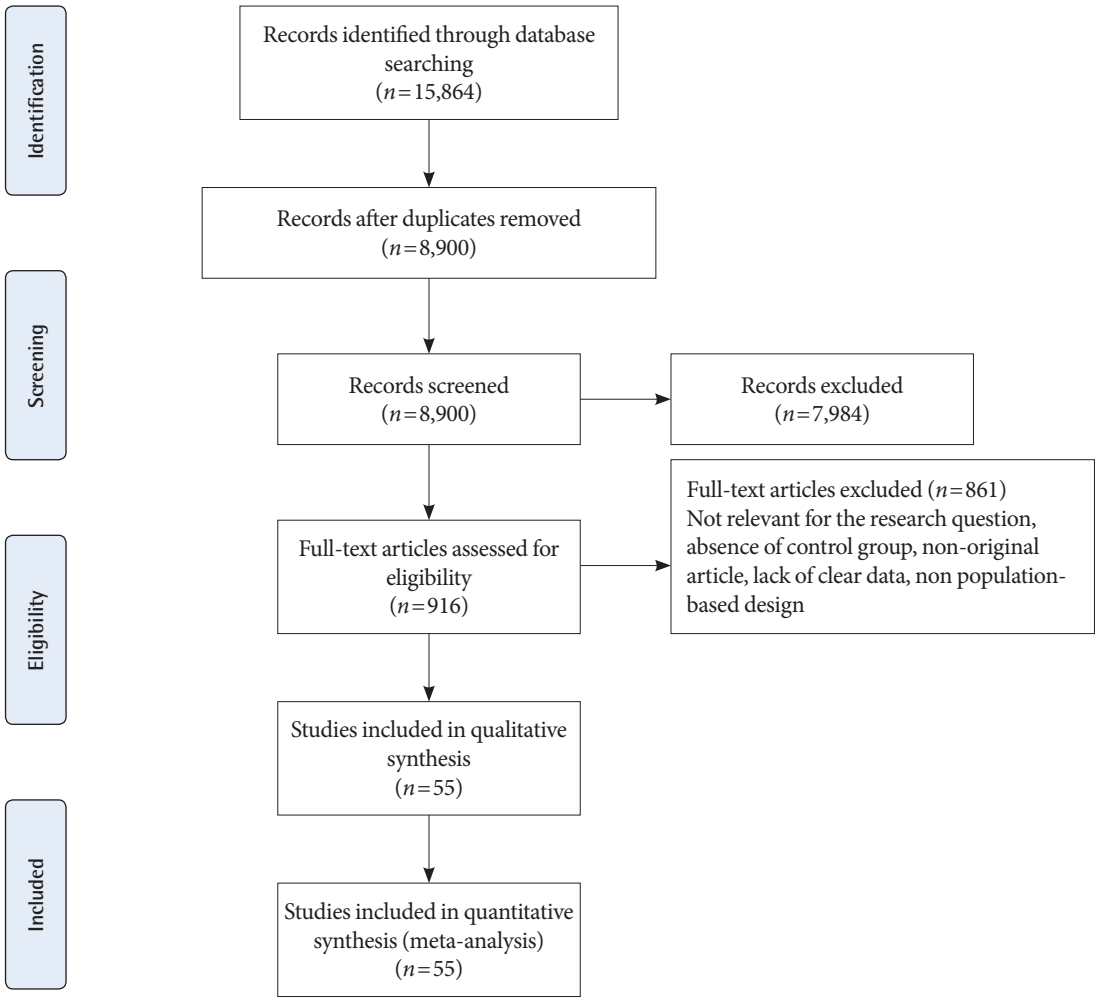


Fig. 1. Flow diagram of the search strategy and study selection.

cross-sectional studies [66,84]. A total of 19 studies classified as group 1 [32,36,48,49,53,57,59,60,65-67,69-71,77-79,81,85] which used IADPSG criteria; six as group 2 [38,51,67,77,83, 86]; three as group 3 [41,58,74]; two as group 4 [37,67]; 22 as group 5 [33-35,39,42,44,45,47,50,52-54,56,59,61,69,72,73,75, 76,82,85]; eight as group 6 [43,46,54,55,61-63,68], and six as group 7 [40,44,64,65,80,84]. It should be noted that 10 studies used more than one GDM classification [44,53,54,59,61,65,67, 69,77,85]. Seventeen studies were conducted in America including USA [32,34,42,49,54,55,61,68,75,76,85] and Canada [35,45,46,69,72,73]; six in Australia [38,51,83,86], New Zealand [67], and Cook Islands [66]; 13 in Asia including Iran [59,60], China [36,53,71,84], Saudi Arabia [78], India [41,58], Korea [52], Qatar [48], and Japan [44,57]; 18 in Europe, including Italy [81], Sweden [37,80], Ireland [64,70,79], UK [74,77], Israel [43,56,62,82], Croatia [65], Spain [63], Norway [40], and Finland [33,39,47]; and one in Mediterranean countries including Malta, Greece, Serbia, Italy, France, Portugal, Morocco, Tunisia, Algeria, Syria, and Lebanon [50].

Meta-analysis and meta-regression results

Figs. 2-4 and Supplementary Figs. 1-6 present the forest plot of outcome measures obtained from Mantel–Haenszel method. Table 2 shows the overall pooled OR (95% CI) of adverse neonatal outcomes, its heterogeneity and publication bias estimation among various subgroups of GDM diagnosis criteria, compared to non-GDM groups.

Results of meta-analyses showed that, regardless of GDM screening criteria, the risk of adverse neonatal outcomes including LGA (pooled overall OR, 2.02; 95% CI, 1.67 to 2.43), NICU admission (pooled overall OR, 1.68; 95% CI, 1.53 to 1.85), preterm birth (pooled overall OR, 1.41; 95% CI, 1.25 to

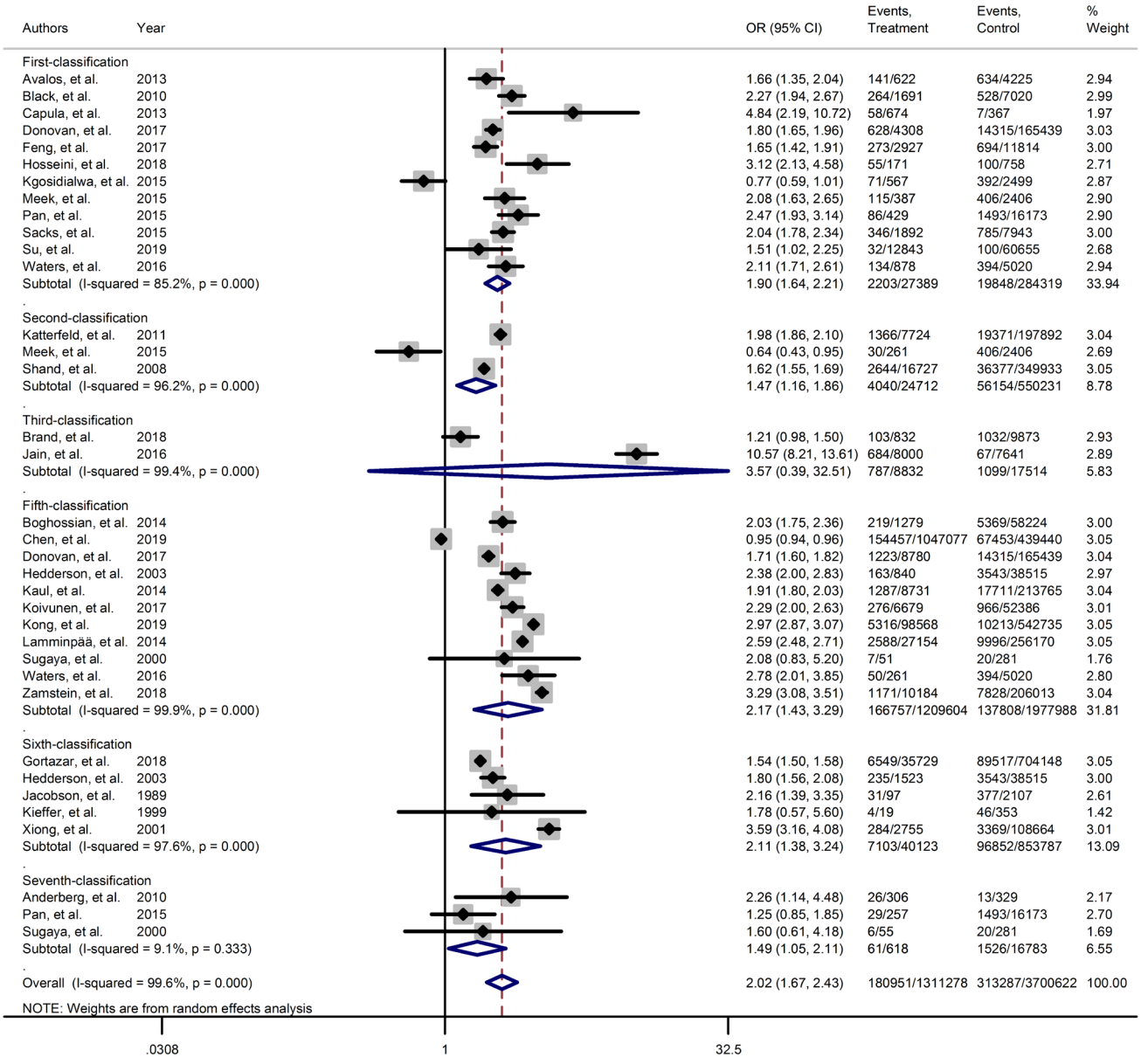


Fig. 2. Forest plot of large for gestational age obtained from Mantel–Haenszel method. Effect size (odds ratio [OR]) and 95% confidence intervals (CIs) for; pooled estimates of effect size are indicated by vertical points of diamonds and 95% CI are represented by horizontal points.

1.60), neonatal hypoglycemia (pooled overall OR, 4.84; 95% CI, 3.24 to 7.25), birth trauma (pooled overall OR, 1.51; 95% CI, 1.24 to 1.82), macrosomia (pooled overall OR, 1.61; 95% CI, 1.43 to 1.82), hyperbilirubinemia (pooled overall OR, 1.50; 95% CI, 1.22 to 1.86), and RDS (pooled overall OR, 1.51; 95% CI, 1.23 to 1.85) significantly increased in women with GDM as compared with the non-GDM group. However, the adverse events of stillbirth was not significantly different between the groups (pooled overall OR, 1.06; 95% CI, 0.78 to 1.44) and the risk of SGA in women with GDM was 0.2 fold lower than in non-GDM (pooled overall OR, 0.80; 95% CI, 0.69 to 0.92).

However, the same results were found for subgroup of GDM diagnostic classification analyses. In this respect, the subgroups analyses demonstrated that the risk of adverse neonatal events including LGA, macrosomia, hyperbilirubinemia, NICU admission, neonatal hypoglycemia, preterm birth, and birth

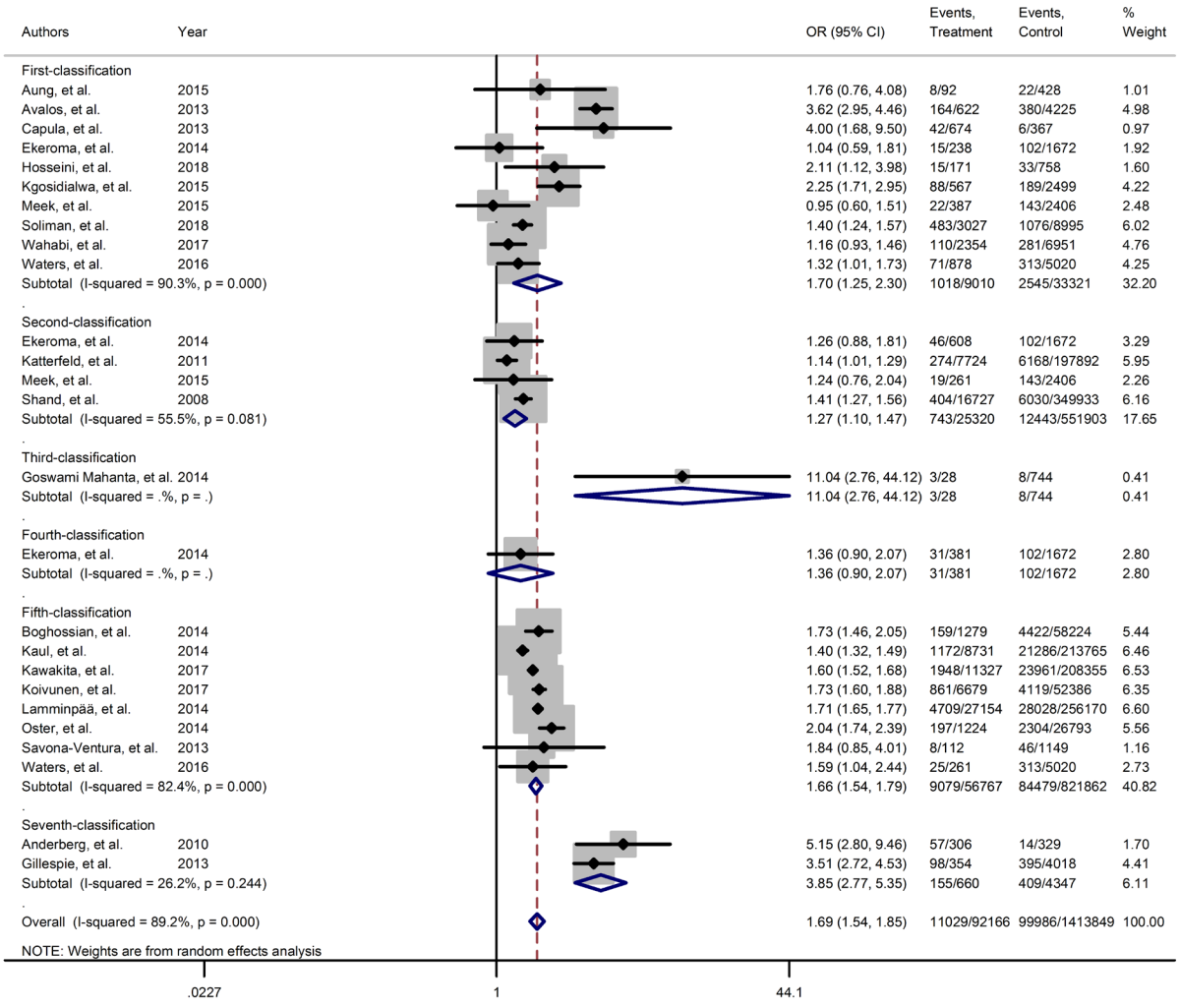


Fig. 3. Forest plot of neonatal intensive care unit obtained from Mantel–Haenszel method. Effect size (odds ratio [OR]) and 95% confidence intervals (CIs) for; pooled estimates of effect size are indicated by vertical points of diamonds and 95% CI are represented by horizontal points.

trauma in women with GDM in all of the GDM diagnostic classification were significantly higher than non-GDM counterparts (Table 2). For example subgroup analyses in IADPSG classification showed that the risk of LGA (pooled OR, 1.90; 95% CI, 1.64 to 2.21), macrosomia (pooled OR, 1.59; 95% CI, 1.35 to 1.89), hyperbilirubinemia (pooled OR, 1.17; 95% CI, 1.07 to 1.28), RDS (pooled OR, 1.60; 95% CI, 1.18 to 2.15), NICU admission (pooled OR, 1.70; 95% CI, 1.25 to 2.30), neonatal hypoglycemia (pooled OR, 4.16; 95% CI, 2.42 to 7.17), preterm birth (pooled OR, 1.32; 95% CI, 1.18 to 1.48), and birth trauma (pooled OR, 1.48; 95% CI, 1.24 to 1.69) in women with GDM were significantly higher than non-GDM population. As well, no significant results were found in the risk of stillbirth (pooled OR, 0.71; 95% CI, 0.46 to 1.09) and also the risk of SGA (pooled OR, 0.77; 95% CI, 0.66 to 0.91) was significantly lower than in non-GDM counterparts.

However, the results of meta-regression revealed that the magnitude of the risk of those adverse neonatal outcomes in the IADPSG criteria, as the strictest one, was similar to other classification (Supplementary Fig. 7).

Publication bias and risk of bias

The results of Begg's test showed that there were no substantial publication bias among various outcomes, except for outcome of LGA, which was corrected by trim and fill method of correction (Table 2). Majority of included studies were judged to

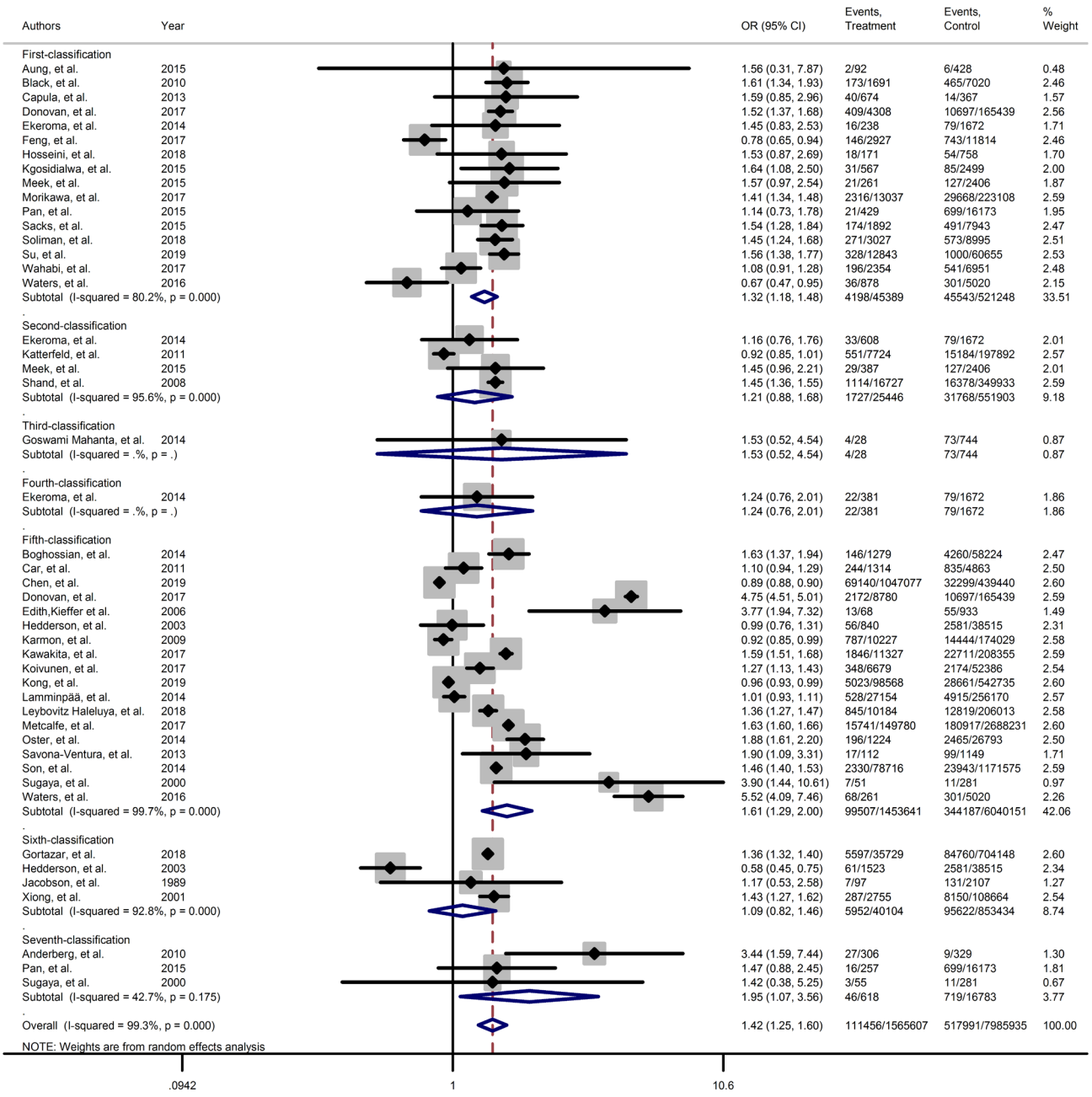


Fig. 4. Forest plot of preterm obtained from Mantel–Haenszel method. Effect size (odds ratio [OR]) and 95% confidence intervals (CIs) for; pooled estimates of effect size are indicated by vertical points of diamonds and 95% CI are represented by horizontal points.

be at low risk of bias for evaluated domains (Supplementary Figs. 8 and 9). Majority of cross-sectional studies had a low or probably low risk of bias in the development of outcome of interest in case and controls, selection of cases and controls and also in assessment of exposure domains. However, half of them had probably high risk of bias in control of prognostic variables. Moreover, cohort studies were judged to have at low risk of bias for selection of exposed and non-exposed cohorts, presence of outcome of interest at start of study, outcome assessment, assessment of prognostic factors and adequacy of follow-up of cohorts; however, one-fifth of them of them had probable high risk of bias in assessment of exposure and 15%

Table 2. Results of heterogeneity and publication bias estimation and subgroup meta-analysis for risk adverse neonatal outcome among women with gestational diabetes based on various GDM screening strategy group

Outcomes[a]	GDM classification	Publication bias[b]	Heterogeneity		Sample size		OR (95% CI)	*P* value from meta-regression
			I^2, %	*P* value	GDM	Control		
SGA	1	0.881	39.2	0.130	21,877	243,938	0.77 (0.66–0.91)	Ref
	2	0.317	81.4	0.021	16,988	352,339	1.03 (0.65–1.62)	0.256
	5	0.458	99.4	0.001	1,204,066	1,912,549	0.74 (0.55–0.98)	0.236
	6	0.174	0	0.988	40,026	851,680	0.90 (0.87–0.94)	0.874
	Overall	0.059	98.4	0.001	1,284,095	3,370,708	0.80 (0.70–0.93)[c]	-
LGA	1	0.583	85.2	0.001	27,389	284,319	1.90 (1.64–2.21)	Ref
	2	0.602	96.2	0.001	24,712	550,231	1.47 (1.16–1.86)	0.212
	3	0.317	99.4	0.001	8,832	17,514	3.57 (0.39–32.51)	0.365
	5	0.024[c]	99.9	0.001	1,209,604	1,977,988	2.17 (1.43–3.29)	0.845
	6	0.624	96.7	0.001	40,123	853,787	2.11 (1.38–3.24)	0.365
	7	0.602	9.1	0.335	618	16,783	1.49 (1.05, 2.11)	0.965
	Overall	0.000[c]	99.6	0.001	1,311,278	3,700,622	2.02[b] (1.67–2.43)[c]	-
Macrosomia	1	0.464	86	0.001	29,846	315,420	1.59 (1.35–1.89)	Ref
	2	1.000	76.3	0.006	1,058	4,480	0.72 (0.39–1.33)	0.251
	4	0.317	87.5	0.005	497	13,938	1.36 (0.42–4.37)	0.452
	5	1.000	98.4	0.001	54,351	1,090,454	1.94 (1.51 –2.49)	0.369
	6	0.317	90.7	0.001	40,641	856,341	1.57 (1.28 –1.93)	0.569
	7	0.497	73.9	0.009	2,779	58,909	1.36 (0.97–1.90)	0.854
	Overall	0.445	96.2	0.009	129,200	2,340,286	1.62 (1.43–1.82)[c]	-
Hyperbilirubinemia	1	0.188	14.4	0.322	8,333	30,103	1.17 (1.07–1.28)	Ref
	5	1.000	64.4	0.024	8,382	117,060	1.37 (1.12–1.68)	0.526
	6	0.317	93.4	0.001	539	2,931	2.82 (0.63–12.60)	0.687
	Overall	0.458	90.7	0.001	25,309	158,016	1.51 (1.22–1.86)[c]	-
Stillbirth	1	0.458	54.7	0.031	24,625	425,869	0.71 (0.46–1.09)	Ref
	2	0.602	0	0.381	17,596	354,011	1.13 (0.87–1.48)	0.251
	3	0.317	0	0.999	8,028	8,385	2.35 (1.87, 2.97)	0.236
	4	0.317	93.2	0.001	497	13,938	2.19 (0.00–1,045.39)	0.028[c]
	5	0.805	71.7	0.001	33,075	638,980	1.14 (0.77–1.69)	0.017[c]
	6	0.602	0	0.634	3,044	114,961	0.94 (0.49–1.79)	0.258
	Overall	0.662	79.3	0.001	87,122	1,572,317	1.07 (0.79–1.45)	-
RDS	1	0.117	0	0.396	2,737	9,068	1.60 (1.18–2.15)	Ref
	5	0.327	55.2	0.063	19,448	320,395	1.41 (1.10–1.80)	0.256
	6	0.317	68.8	0.073	539	2,931	1.60 (0.24–10.49)	0.854
	Overall	0.586	44	0.057	22,779	332,675	1.51 (1.23–1.85)[c]	-
NICU admission	1	0.245	90.3	0.001	9,010	33,321	1.70 (1.25–2.30)	Ref
	2	0.497	55.5	0.081	25,320	551,903	1.27 (1.10 –1.47)	0.258
	5	0.621	82.4	0.001	56,767	821,862	1.66 (1.54–1.79)	0.369
	Overall	0.280	89.2	0.001	92,166	1,413,849	1.69 (1.54–1.85)[c]	-
Neonatal hypoglycemia	1	1.000	75.2	0.003	7,038	24,824	4.16 (2.42–7.17)	Ref
	5	1.000	97.9	0.001	8,270	115,911	2.78 (0.57–13.50)	0.523
	Overall	0.186	97.4	0.001	32,132	492,775	4.84 (3.24–7.25)[c]	-

(*Continued to the next page*)

Table 2. Continued

Outcomes[a]	GDM classification	Publication bias[b]	Heterogeneity		Sample size		OR (95% CI)	*P* value from meta-regression
			I^2, %	*P* value	GDM	Control		
Preterm	1	0.105	80.2	0.001	45,839	521,248	1.32 (1.18–1.48)	Ref
	2	1.000	95.6	0.001	25,446	551,903	1.21 (0.88–1.68)	0.254
	5	0.910	99.7	0.001	1,453,641	6,040,151	1.61 (1.29–2.00)	0.165
	6	0.174	92.8	0.001	40,104	853,434	1.09 (0.82–1.46)	0.895
	7	0.602	42.7	0.175	618	16,783	1.95 (1.07–3.56)	0.207
	Overall	0.166	99.3	0.001	1,565,607	7,985,935	1.42 (1.25–1.60)[c]	-
Birth trauma	1	1.000	0	0.443	7,178	24,780	1.45 (1.24–1.69)	Ref
	2	0.317	89.6	0.002	24,451	547,825	1.28 (1.00–1.66)	0.257
	5	0.652	78.2	0.001	124,428	1,718,553	1.57 (1.24–2.00)	0.584
	6	0.317	72	0.059	539	2,931	1.34 (0.10–17.81)	0.985
	Overall	0.368	88.1	0.001	156,596	2,294,089	1.51 (1.25–1.83)[c]	-

GDM, gestational diabetes mellitus; OR, odds ratio; CI, confidence interval; SGA, small for gestational age; LGA, large for gestational age; RDS, respiratory distress syndrome; NICU, neonatal intensive care unit.
[a]All subgroups analyses did not performed due to lack of available data, [b]Obtained from trim and fill method, [c]Statistically significant level $P<0.05$.

of them had high risk of bias controlling prognostic variables.

DISCUSSION

In this meta-analysis of observational studies, we evaluated the impact of several diagnostic criteria for GDM on the risk of adverse-neonatal-outcomes. Briefly, the results showed that neonates of women with GDM have higher risk of adverse outcomes of LGA macrosomia, hyperbilirubinemia, RDS, neonatal hypoglycemia, NICU admission, preterm birth and birth trauma than the neonates of women without GDM group; however, the magnitude of these adverse neonatal outcomes were not significantly varies by different diagnostic criteria for GDM.

During pregnancy failure to adapt with physiological changes of pregnancy as a result of the dysfunction in pancreatic β-cell, and developing the GDM considered as a threatening factor for maternal and child health [87]. It is well established that the risk of adverse neonatal outcomes are increased among women with GDM [88].

There are many guidelines which provide various recommendations for screening and diagnosis criteria of GDM [89,90], but debate about the screening and diagnosis for GDM still continue in the literature. Different approaches identify different feto-maternal and neonatal risks leading to variation in prevalence and pregnancy related outcomes [3,32-86,91]. However stringent criteria of IADPSG, that is accepted by many organizations, led to increase of GDM cases [3]. However, there are limited evidence to support the IADPSG criteria to prove clinically significant improvements in maternal and neonatal outcomes. The main purpose of these struggles is to find a practical strategy with minimum costs, adverse maternal-fetal outcomes and maximum availability especially in low health resources countries. The results of this systemic review and meta-analysis confirmed the previous findings about increased risk of adverse neonatal outcomes among women with GDM. In addition, it revealed that the magnitude of those increased risk are similar in various GDM diagnostic criteria.

There are extensive discussions regarding the cost-effectiveness of different GDM diagnosis criteria in the literature [92-95]. Considering that those increased cased without any improvement in pregnancy outcomes potentially may lead to over medicalization of pregnant women [96,97] and therefore increased health costs, and decreased physical, psychological, social, and other aspects of quality of life in pregnant women [98].

In line with the findings of current meta-analysis, in another our recent published meta-analysis with the same classification for GDM screening, we found that magnitude of the risks of adverse maternal outcomes including primary cesarean section, induction of labor, maternal hemorrhage, pregnancy related hypertension, and gestational weight gain are similar all

GDM screening strategies classifications [99]. In agreement, Wendland et al. [88] (2012), in a systematic review study demonstrated that risk of LGA among participants with GDM was higher than non-GDM counterparts in both WHO and the IADPSG criteria, and also the magnitude of this risk was similar in both criteria. Hartling et al. [100] (2014), in their meta-analysis found higher glucose thresholds did not consistently demonstrate greater risk of adverse pregnancy outcomes. Further, Hosseini and Janghorbani [101] (2018) in a meta-analysis reported that women with GDM diagnosed with either the one-step or the two-step approach were at increased risk for selected adverse pregnancy outcomes. The associations with the two-step method were slightly stronger. However, all of these mentioned studies have not compared the various existing criteria and did not provided the majority of neonatal outcomes.

However, in the present study, the risk of stillbirth was not significantly different between the women with GDM and non-GDM groups. Additionally, compared to non-GDM women, the risk of SGA was significantly lower in women with GDM. It may be due to that all of women diagnosed with GDM have been received glucose lowering therapy in order to decrease the feto-maternal adverse outcomes, particularly some sever outcomes such as stillbirth. It is well known that optimal control of maternal blood glucose could strongly decrease risk of still birth [102]. Also it should be noted that stillbirth rates vary based on the management option (insulin/diet) and gestational week, but due to the limitations of the data we cannot adjust the mentioned factors. In addition it is well documented that intensive therapy for GDM may affect the fetal growth and increase the SGA rate [103,104]. Moreover, vasculopathy plays a role in increased risk of SGA in GDM suffering women [105-107].

This review has certain strengths and limitations. Population-based design of included studies with high quality, large sample size of GDM, and non-GDM participants from different countries, estimation of the pooled risk of several neonatal outcomes in different subgroups of GDM classifications let us to present reliable evidence. In addition, Subgroup analysis and assessed multiple available GDM screening and diagnostic criteria were considered as the strength of our study. However, our meta-analysis has limitations, such as the presence of significant heterogeneity in some subgroup analyses, only studies published in English included, and did not investigate grey literature. In addition, due to lack of data, we could not perform some subgroup analysis. Additionally, different definitions of outcome measures across included studies may impose potential limitations in this meta-analysis. There is a need for future meta-analysis and observational studies about the long term effect of GDM from childhood into adulthood GDM based on the different classifications of GDM diagnosis criteria.

In conclusion, our results showed that the risk of adverse neonatal outcome increased among women with GDM, but the magnitude of risk was not different among those women who were diagnosed through more or less intensive strategies. These findings may help health-care-providers and policy makers to select the most cost-effective approach for the screening of GDM among pregnant women.

SUPPLEMENTARY MATERIALS

Supplementary materials related to this article can be found online at https://doi.org/10.4093/dmj.2021.0178.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

AUTHOR CONTRIBUTIONS

Conception or design: F.R.T., S.B.G.
Acquisition, analysis, or interpretation of data: F.R.T., M.S.G.N., R.B.Y., S.B.G.
Drafting the work or revising: F.R.T., M.S.G.N., R.B.Y., S.B.G.
Final approval of the manuscript: F.R.T., M.S.G.N., R.B.Y., S.B.G.

ORCID

Fahimeh Ramezani Tehrani
https://orcid.org/0000-0002-4609-065X
Samira Behboudi-Gandevani
https://orcid.org/0000-0003-3526-640X

FUNDING

This research project was funded by the Shahid Beheshti University of Medical Sciences. Also, Nord University, Bodø, Norway covered the article processing charges.

ACKNOWLEDGMENTS

The authors wish to acknowledge Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

REFERENCES

1. Kampmann U, Madsen LR, Skajaa GO, Iversen DS, Moeller N, Ovesen P. Gestational diabetes: a clinical update. World J Diabetes 2015;6:1065-72.
2. American Diabetes Association. 2. Classification and diagnosis of diabetes: standards of medical care in diabetes-2020. Diabetes Care 2020;43(Suppl 1):S14-31.
3. Behboudi-Gandevani S, Amiri M, Bidhendi Yarandi R, Ramezani Tehrani F. The impact of diagnostic criteria for gestational diabetes on its prevalence: a systematic review and meta-analysis. Diabetol Metab Syndr 2019;11:11.
4. Baz B, Riveline JP, Gautier JF. Endocrinology of pregnancy: gestational diabetes mellitus: definition, aetiological and clinical aspects. Eur J Endocrinol 2016;174:R43-51.
5. Li Y, Ren X, He L, Li J, Zhang S, Chen W. Maternal age and the risk of gestational diabetes mellitus: a systematic review and meta-analysis of over 120 million participants. Diabetes Res Clin Pract 2020;162:108044.
6. Karacam Z, Celik D. The prevalence and risk factors of gestational diabetes mellitus in Turkey: a systematic review and meta-analysis. J Matern Fetal Neonatal Med 2021;34:1331-41.
7. Shou C, Wei YM, Wang C, Yang HX. Updates in long-term maternal and fetal adverse effects of gestational diabetes mellitus. Matern Fetal Med 2019;1:91-4.
8. Nijs H, Benhalima K. Gestational diabetes mellitus and the long-term risk for glucose intolerance and overweight in the offspring: a narrative review. J Clin Med 2020;9:599.
9. Muche AA, Olayemi OO, Gete YK. Effects of gestational diabetes mellitus on risk of adverse maternal outcomes: a prospective cohort study in Northwest Ethiopia. BMC Pregnancy Childbirth 2020;20:73.
10. Stewart A, Malhotra A. Gestational diabetes and the neonate: challenges and solutions. Res Rep Neonatol 2015;5:31-9.
11. Bhavadharini B, Uma R, Saravanan P, Mohan V. Screening and diagnosis of gestational diabetes mellitus: relevance to low and middle income countries. Clin Diabetes Endocrinol 2016;2:13.
12. International Association of Diabetes and Pregnancy Study Groups Consensus Panel, Metzger BE, Gabbe SG, Persson B, Buchanan TA, Catalano PA, et al. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care 2010;33:676-82.
13. HAPO Study Cooperative Research Group, Metzger BE, Lowe LP, Dyer AR, Trimble ER, Chaovarindr U, et al. Hyperglycemia and adverse pregnancy outcomes. N Engl J Med 2008;358:1991-2002.
14. Hod M, Kapur A, Sacks DA, Hadar E, Agarwal M, Di Renzo GC, et al. The International Federation of Gynecology and Obstetrics (FIGO) Initiative on gestational diabetes mellitus: a pragmatic guide for diagnosis, management, and care. Int J Gynaecol Obstet 2015;131 Suppl 3:S173-211.
15. Diagnostic criteria and classification of hyperglycaemia first detected in pregnancy: a World Health Organization Guideline. Diabetes Res Clin Pract 2014;103:341-63.
16. Agarwal MM. Gestational diabetes mellitus: an update on the current international diagnostic criteria. World J Diabetes 2015;6:782-91.
17. Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, Thompson D, Berger H, Feig D, Gagnon R, Kader T, et al. Diabetes and pregnancy. Can J Diabetes 2013;37 Suppl 1:S168-83.
18. Buckley BS, Harreiter J, Damm P, Corcoy R, Chico A, Simmons D, et al. Gestational diabetes mellitus in Europe: prevalence, current screening practice and barriers to screening: a review. Diabet Med 2012;29:844-54.
19. Nankervis A, McIntyre HD, Moses R, Ross GP, Callaway L, Porter C, et al: ADIPS consensus guidelines for the testing and diagnosis of hyperglycaemia in pregnancy in Australia and New Zealand. Available from: Available from: http://adips.org/downloads/2014ADIPSGDMGuidelinesV18.11.2014_000.pdf (cited 2021 Dec 4).
20. New Zealand Ministry of Health. Screening, diagnosis and management of gestational diabetes in New Zealand: a clinical practice guideline. Wellington: Ministry of Health; 2014 [cited 2021 Dec 4]. Available from: http://www.health.govt.nz/publication/screening-diagnosis-and-management-gestational-diabetes-new-zealand-clinical-practice-guideline.
21. Japan Diabetes Society: Evidence-based practice guideline for the treatment for diabetes in Japan 2013. Available from: http://www.jds.or.jp/modules/en/index.php?content_id=44 (cited 2021 Dec 4).
22. Trujillo J, Vigo A, Reichelt A, Duncan BB, Schmidt MI. Fast-

ing plasma glucose to avoid a full OGTT in the diagnosis of gestational diabetes. Diabetes Res Clin Pract 2014;105:322-6.
23. International Diabetes Federation: Global guidelines for type 2 diabetes. Available from: https://www.idf.org/e-library/guidelines/79-global-guideline-for-type-2-diabetes (cited 2021 Dec 4).
24. Hur KY. New diagnostic criteria for gestational diabetes mellitus and pregnancy outcomes in Korea. Diabetes Metab J 2019;43:763-5.
25. ACOG Practice Bulletin no. 190: gestational diabetes mellitus. Obstet Gynecol 2018;131:e49-64.
26. Landon MB. Changing the diagnostic criteria for gestational diabetes mellitus? Obstet Gynecol 2016;127:3-6.
27. Vandorsten JP, Dodson WC, Espeland MA, Grobman WA, Guise JM, Mercer BM, et al. NIH consensus development conference: diagnosing gestational diabetes mellitus. NIH Consens State Sci Statements 2013;29:1-31.
28. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8:336-41.
29. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions (version 5.1. 0). New York: Wiley; 2011. Chapter 9, Analysing data and undertaking meta-analyses [cited 2021 Dec 4]. Available from: https://training.cochrane.org/handbook/archive/v5.1.
30. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. New York: John Wiley & Sons; 2011.
31. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56:455-63.
32. Black MH, Sacks DA, Xiang AH, Lawrence JM. Clinical outcomes of pregnancies complicated by mild gestational diabetes mellitus differ by combinations of abnormal oral glucose tolerance test values. Diabetes Care 2010;33:2524-30.
33. Lamminpaa R, Vehvilainen-Julkunen K, Gissler M, Selander T, Heinonen S. Pregnancy outcomes in women aged 35 years or older with gestational diabetes: a registry-based study in Finland. J Matern Fetal Neonatal Med 2016;29:55-9.
34. Carr DB, Newton KM, Utzschneider KM, Faulenbach MV, Kahn SE, Easterling TR, et al. Gestational diabetes or lesser degrees of glucose intolerance and risk of preeclampsia. Hypertens Pregnancy 2011;30:153-63.
35. Metcalfe A, Sabr Y, Hutcheon JA, Donovan L, Lyons J, Burrows J, et al. Trends in obstetric intervention and pregnancy outcomes of Canadian women with diabetes in pregnancy from 2004 to 2015. J Endocr Soc 2017;1:1540-9.
36. Su WJ, Chen YL, Huang PY, Shi XL, Yan FF, Chen Z, et al. Effects of prepregnancy body mass index, weight gain, and gestational diabetes mellitus on pregnancy outcomes: a population-based study in Xiamen, China, 2011-2018. Ann Nutr Metab 2019;75:31-8.
37. Aberg A, Rydhstroem H, Frid A. Impaired glucose tolerance associated with adverse pregnancy outcome: a population-based study in southern Sweden. Am J Obstet Gynecol 2001;184:77-83.
38. van Hoorn J, Dekker G, Jeffries B. Gestational diabetes versus obesity as risk factors for pregnancy-induced hypertensive disorders and fetal macrosomia. Aust N Z J Obstet Gynaecol 2002;42:35-40.
39. Kong L, Nilsson IA, Gissler M, Lavebratt C. Associations of maternal diabetes and body mass index with offspring birth weight and prematurity. JAMA Pediatr 2019;173:371-8.
40. Sletner L, Jenum AK, Yajnik CS, Morkrid K, Nakstad B, Rognerud-Jensen OH, et al. Fetal growth trajectories in pregnancies of European and South Asian mothers with and without gestational diabetes, a population-based cohort study. PLoS One 2017;12:e0172946.
41. Mahanta TG, Deuri A, Mahanta BN, Bordoloi P, Rasaily R, Mahanta J, et al. Maternal and foetal outcome of gestational diabetes mellitus in a rural block of Assam, India. Clin Epidemiol Glob Health 2014;2:9-15.
42. Kieffer EC, Tabaei BP, Carman WJ, Nolan GH, Guzman JR, Herman WH. The influence of maternal weight and glucose tolerance on infant birthweight in Latino mother-infant pairs. Am J Public Health 2006;96:2201-8.
43. Fraser D, Weitzman S, Leiberman JR, Zmora E, Laron E, Karplus M. Gestational diabetes among Bedouins in southern Israel: comparison of prevalence and neonatal outcomes with the Jewish population. Acta Diabetol 1994;31:78-81.
44. Sugaya A, Sugiyama T, Nagata M, Toyoda N. Comparison of the validity of the criteria for gestational diabetes mellitus by WHO and by the Japan Society of Obstetrics and Gynecology by the outcomes of pregnancy. Diabetes Res Clin Pract 2000;50:57-63.
45. Oster RT, King M, Morrish DW, Mayan MJ, Toth EL. Diabetes in pregnancy among First Nations women in Alberta, Canada: a retrospective analysis. BMC Pregnancy Childbirth 2014;14:136.
46. Xiong X, Saunders LD, Wang FL, Demianczuk NN. Gestational diabetes mellitus: prevalence, risk factors, maternal and

infant outcomes. Int J Gynaecol Obstet 2001;75:221-8.
47. Koivunen S, Torkki A, Bloigu A, Gissler M, Pouta A, Kajantie E, et al. Towards national comprehensive gestational diabetes screening: consequences for neonatal outcome and care. Acta Obstet Gynecol Scand 2017;96:106-13.
48. Soliman A, Salama H, Al Rifai H, De Sanctis V, Al-Obaidly S, Al Qubasi M, et al. The effect of different forms of dysglycemia during pregnancy on maternal and fetal outcomes in treated women and comparison with large cohort studies. Acta Biomed 2018;89(S5):11-21.
49. Sacks DA, Black MH, Li X, Montoro MN, Lawrence JM. Adverse pregnancy outcomes using the International Association of the Diabetes and Pregnancy Study Groups criteria: glycemic thresholds and associated risks. Obstet Gynecol 2015;126:67-73.
50. Savona-Ventura C, Vassallo J, Marre M, Karamanos BG; MGSD-GDM study group. A composite risk assessment model to screen for gestational diabetes mellitus among Mediterranean women. Int J Gynaecol Obstet 2013;120:240-4.
51. von Katterfeld B, Li J, McNamara B, Langridge AT. Maternal and neonatal outcomes associated with gestational diabetes in women from culturally and linguistically diverse backgrounds in Western Australia. Diabet Med 2012;29:372-7.
52. Son KH, Lim NK, Lee JW, Cho MC, Park HY. Comparison of maternal morbidity and medical costs during pregnancy and delivery between patients with gestational diabetes and patients with pre-existing diabetes. Diabet Med 2015;32:477-86.
53. Pan L, Leng J, Liu G, Zhang C, Liu H, Li M, et al. Pregnancy outcomes of Chinese women with gestational diabetes mellitus defined by the IADPSG's but not by the 1999 WHO's criteria. Clin Endocrinol (Oxf) 2015;83:684-93.
54. Schwartz ML, Ray WN, Lubarsky SL. The diagnosis and classification of gestational diabetes mellitus: is it time to change our tune? Am J Obstet Gynecol 1999;180(6 Pt 1):1560-71.
55. Jacobson JD, Cousins L. A population-based study of maternal and perinatal outcome in patients with gestational diabetes. Am J Obstet Gynecol 1989;161:981-6.
56. Leybovitz-Haleluya N, Wainstock T, Landau D, Sheiner E. Maternal gestational diabetes mellitus and the risk of subsequent pediatric cardiovascular diseases of the offspring: a population-based cohort study with up to 18 years of follow up. Acta Diabetol 2018;55:1037-42.
57. Morikawa M, Sugiyama T, Sagawa N, Hiramatsu Y, Ishikawa H, Hamada H, et al. Perinatal mortality in Japanese women diagnosed with gestational diabetes mellitus and diabetes mellitus. J Obstet Gynaecol Res 2017;43:1700-7.
58. Jain R, Davey S, Davey A, Raghav SK, Singh JV. Can the management of blood sugar levels in gestational diabetes mellitus cases be an indicator of maternal and fetal outcomes? The results of a prospective cohort study from India. J Family Community Med 2016;23:94-9.
59. Hosseini E, Janghorbani M, Aminorroaya A. Incidence, risk factors, and pregnancy outcomes of gestational diabetes mellitus using one-step versus two-step diagnostic approaches: a population-based cohort study in Isfahan, Iran. Diabetes Res Clin Pract 2018;140:288-94.
60. Hosseini E, Janghorbani M, Shahshahan Z. Comparison of risk factors and pregnancy outcomes of gestational diabetes mellitus diagnosed during early and late pregnancy. Midwifery 2018;66:64-9.
61. Hedderson MM, Ferrara A, Sacks DA. Gestational diabetes mellitus and lesser degrees of pregnancy hyperglycemia: association with increased risk of spontaneous preterm birth. Obstet Gynecol 2003;102:850-6.
62. Zamstein O, Sheiner E, Wainstock T, Landau D, Walfisch A. Maternal gestational diabetes and long-term respiratory related hospitalizations of the offspring. Diabetes Res Clin Pract 2018;140:200-7.
63. Gortazar L, Flores-Le Roux JA, Benaiges D, Sarsanedas E, Paya A, Mane L, et al. Trends in prevalence of gestational diabetes and perinatal outcomes in Catalonia, Spain, 2006 to 2015: the Diagestcat Study. Diabetes Metab Res Rev 2019;35:e3151.
64. Gillespie P, Cullinan J, O'Neill C, Dunne F; ATLANTIC DIP Collaborators. Modeling the independent effects of gestational diabetes mellitus on maternity care and costs. Diabetes Care 2013;36:1111-6.
65. Erjavec K, Poljicanin T, Matijevic R. Impact of the implementation of new WHO diagnostic criteria for gestational diabetes mellitus on prevalence and perinatal outcomes: a population-based study. J Pregnancy 2016;2016:2670912.
66. Aung YY, Sowter M, Kenealy T, Herman J, Ekeroma A. Gestational diabetes mellitus screening, management and outcomes in the Cook Islands. N Z Med J 2015;128:21-8.
67. Ekeroma AJ, Chandran GS, McCowan L, Ansell D, Eagleton C, Kenealy T. Impact of using the International Association of Diabetes and Pregnancy Study Groups criteria in South Auckland: prevalence, interventions and outcomes. Aust N Z J Obstet Gynaecol 2015;55:34-41.
68. Kieffer EC, Nolan GH, Carman WJ, Sanborn CZ, Guzman R, Ventura A. Glucose tolerance during pregnancy and birth

weight in a Hispanic population. Obstet Gynecol 1999;94(5 Pt 1):741-6.
69. Donovan LE, Edwards AL, Savu A, Butalia S, Ryan EA, Johnson JA, et al. Population-level outcomes with a 2-step approach for gestational diabetes screening and diagnosis. Can J Diabetes 2017;41:596-602.
70. Kgosidialwa O, Egan AM, Carmody L, Kirwan B, Gunning P, Dunne FP. Treatment with diet and exercise for women with gestational diabetes mellitus diagnosed using IADPSG criteria. J Clin Endocrinol Metab 2015;100:4629-36.
71. Feng H, Zhu WW, Yang HX, Wei YM, Wang C, Su RN, et al. Relationship between oral glucose tolerance test characteristics and adverse pregnancy outcomes among women with gestational diabetes mellitus. Chin Med J (Engl) 2017;130:1012-8.
72. Chen L, Wang WJ, Auger N, Xiao L, Torrie J, McHugh NG, et al. Diabetes in pregnancy in associations with perinatal and postneonatal mortality in First Nations and non-Indigenous populations in Quebec, Canada: population-based linked birth cohort study. BMJ Open 2019;9:e025084.
73. Kaul P, Savu A, Nerenberg KA, Donovan LE, Chik CL, Ryan EA, et al. Impact of gestational diabetes mellitus and high maternal weight on the development of diabetes, hypertension and cardiovascular disease: a population-level analysis. Diabet Med 2015;32:164-73.
74. Brand JS, West J, Tuffnell D, Bird PK, Wright J, Tilling K, et al. Gestational diabetes and ultrasound-assessed fetal growth in South Asian and White European women: findings from a prospective pregnancy cohort. BMC Med 2018;16:203.
75. Kawakita T, Bowers K, Hazrati S, Zhang C, Grewal J, Chen Z, et al. Increased neonatal respiratory morbidity associated with gestational and pregestational diabetes: a retrospective study. Am J Perinatol 2017;34:1160-8.
76. Boghossian NS, Yeung E, Albert PS, Mendola P, Laughon SK, Hinkle S, et al. Changes in diabetes status between pregnancies and impact on subsequent newborn outcomes. Am J Obstet Gynecol 2014; 210: 431.e1-14.
77. Meek CL, Lewis HB, Patient C, Murphy HR, Simmons D. Diagnosis of gestational diabetes mellitus: falling through the net. Diabetologia 2015;58:2003-12.
78. Wahabi H, Fayed A, Esmaeil S, Mamdouh H, Kotb R. Prevalence and complications of pregestational and gestational diabetes in Saudi women: analysis from Riyadh Mother and Baby cohort study (RAHMA). Biomed Res Int 2017;2017:6878263.
79. Avalos GE, Owens LA, Dunne F; ATLANTIC DIP Collaborators. Applying current screening tools for gestational diabetes mellitus to a European population: is it time for change? Diabetes Care 2013;36:3040-4.
80. Anderberg E, Kallen K, Berntorp K. The impact of gestational diabetes mellitus on pregnancy outcome comparing different cut-off criteria for abnormal glucose tolerance. Acta Obstet Gynecol Scand 2010;89:1532-7.
81. Capula C, Chiefari E, Vero A, Arcidiacono B, Iiritano S, Puccio L, et al. Gestational diabetes mellitus: screening and outcomes in southern Italian pregnant women. ISRN Endocrinol 2013;2013:387495.
82. Karmon A, Levy A, Holcberg G, Wiznitzer A, Mazor M, Sheiner E. Decreased perinatal mortality among women with diet-controlled gestational diabetes mellitus. Int J Gynaecol Obstet 2009;104:199-202.
83. Shand AW, Bell JC, McElduff A, Morris J, Roberts CL. Outcomes of pregnancies in women with pre-gestational diabetes mellitus and gestational diabetes mellitus: a population-based study in New South Wales, Australia, 1998-2002. Diabet Med 2008;25:708-15.
84. Gu Y, Lu J, Li W, Liu H, Wang L, Leng J, et al. Joint associations of maternal gestational diabetes and hypertensive disorders of pregnancy with overweight in offspring. Front Endocrinol (Lausanne) 2019;10:645.
85. Waters TP, Dyer AR, Scholtens DM, Dooley SL, Herer E, Lowe LP, et al. Maternal and neonatal morbidity for women who would be added to the diagnosis of GDM using IADPSG criteria: a secondary analysis of the hyperglycemia and adverse pregnancy outcome study. Diabetes Care 2016;39:2204-10.
86. Moses RG, Griffiths RD. Can a diagnosis of gestational diabetes be an advantage to the outcome of pregnancy? J Soc Gynecol Investig 1995;2:523-5.
87. Alejandro EU, Mamerto TP, Chung G, Villavieja A, Gaus NL, Morgan E, et al. Gestational diabetes mellitus: a harbinger of the vicious cycle of diabetes. Int J Mol Sci 2020;21:5003.
88. Wendland EM, Torloni MR, Falavigna M, Trujillo J, Dode MA, Campos MA, et al. Gestational diabetes and pregnancy outcomes: a systematic review of the World Health Organization (WHO) and the International Association of Diabetes in Pregnancy Study Groups (IADPSG) diagnostic criteria. BMC Pregnancy Childbirth 2012;12:23.
89. Li-Zhen L, Yun X, Xiao-Dong Z, Shu-Bin H, Zi-Lian W, Adrian Sandra D, et al. Evaluation of guidelines on the screening and diagnosis of gestational diabetes mellitus: systematic review. BMJ Open 2019;9:e023014.
90. Hong S, Lee SM, Kwak SH, Kim BJ, Koo JN, Oh IH, et al. A

comparison of predictive performances between old versus new criteria in a risk-based screening strategy for gestational diabetes mellitus. Diabetes Metab J 2020;44:726-36.

91. Kim MH, Kwak SH, Kim SH, Hong JS, Chung HR, Choi SH, et al. Pregnancy outcomes of women additionally diagnosed as gestational diabetes by the International Association of the Diabetes and Pregnancy Study Groups Criteria. Diabetes Metab J 2019;43:766-75.
92. Jacklin PB, Maresh MJ, Patterson CC, Stanley KP, Dornhorst A, Burman-Roy S, et al. A cost-effectiveness comparison of the NICE 2015 and WHO 2013 diagnostic criteria for women with gestational diabetes with and without risk factors. BMJ Open 2017;7:e016621.
93. Meltzer SJ, Snyder J, Penrod JR, Nudi M, Morin L. Gestational diabetes mellitus screening and diagnosis: a prospective randomised controlled trial comparing costs of one-step and two-step methods. BJOG 2010;117:407-15.
94. Chen PY, Finkelstein EA, Ng MJ, Yap F, Yeo GS, Rajadurai VS, et al. Incremental cost-effectiveness analysis of gestational diabetes mellitus screening strategies in Singapore. Asia Pac J Public Health 2016;28:15-25.
95. Mission JF, Ohno MS, Cheng YW, Caughey AB. Gestational diabetes screening with the new IADPSG guidelines: a cost-effectiveness analysis. Am J Obstet Gynecol 2012;207:326.
96. Gupta Y, Kalra B, Baruah MP, Singla R, Kalra S. Updated guidelines on screening for gestational diabetes. Int J Womens Health 2015;7:539-50.
97. Visser GH, de Valk HW. Is the evidence strong enough to change the diagnostic criteria for gestational diabetes now? Am J Obstet Gynecol 2013;208:260-4.
98. Marchetti D, Carrozzino D, Fraticelli F, Fulcheri M, Vitacolonna E. Quality of life in women with gestational diabetes mellitus: a systematic review. J Diabetes Res 2017;2017:7058082.
99. Ramezani Tehrani F, Naz M, Yarandi RB, Behboudi-Gandevani S. The impact of diagnostic criteria for gestational diabetes mellitus on adverse maternal outcomes: a systematic review and meta-analysis. J Clin Med 2021;10:666.
100. Hartling L, Dryden DM, Guthrie A, Muise M, Vandermeer B, Donovan L. Diagnostic thresholds for gestational diabetes and their impact on pregnancy outcomes: a systematic review. Diabet Med 2014;31:319-31.
101. Hosseini E, Janghorbani M. Systematic review and meta-analysis of diagnosing gestational diabetes mellitus with one-step or two-step approaches and associations with adverse pregnancy outcomes. Int J Gynaecol Obstet 2018;143:137-44.
102. Mackin ST, Nelson SM, Wild SH, Colhoun HM, Wood R, Lindsay RS, et al. Factors associated with stillbirth in women with diabetes. Diabetologia 2019;62:1938-47.
103. Kjos SL, Schaefer-Graf UM. Modified therapy for gestational diabetes using high-risk and low-risk fetal abdominal circumference growth to select strict versus relaxed maternal glycemic targets. Diabetes Care 2007;30 Suppl 2:S200-5.
104. Langer O, Levy J, Brustman L, Anyaegbunam A, Merkatz R, Divon M. Glycemic control in gestational diabetes mellitus: how tight is tight enough. Small for gestational age versus large for gestational age? Am J Obstet Gynecol 1989;161:646-53.
105. Huynh J, Yamada J, Beauharnais C, Wenger JB, Thadhani RI, Wexler D, et al. Type 1, type 2 and gestational diabetes mellitus differentially impact placental pathologic characteristics of uteroplacental malperfusion. Placenta 2015;36:1161-6.
106. McCowan L, Horgan RP. Risk factors for small for gestational age infants. Best Pract Res Clin Obstet Gynaecol 2009;23:779-93.
107. Pontes IE, Afra KF, Silva JR Jr, Borges PS, Clough GF, Alves JG. Microvascular reactivity in women with gestational diabetes mellitus studied during pregnancy. Diabetol Metab Syndr 2015;7:27.

Supplementary Table 1. Summary of available studies included in the present meta-analysis

Study	Location	GDM diagnostic criteria	GDM characteristics	Non-GDM characteristics	Neonatal outcome in women with vs. without GDM, %
Capula et al. (2013) [81]	Italy	IADPSG	*n*=171, Age: 30.8 (3.2), BMI: 22.8 (1.9)	*n*=367, Age: 29.3 (3.5), BMI: 21.4 (2.0)	Dystocia: 0 vs. 0.3; Admission to NICU: 6.4 vs. 1.6, RDS: 1.8 vs. 0.5; Macrosomia: 1.2 vs. 1.6; LGA: 8.8 vs. 1.9; SGA: 2.9 vs. 1.6; Hypoglycaemia: 0.6 vs. 0; Hyperbilirubinemia: 2.4 vs. 0.8; Bone fracture: 1.2 vs.0; Preterm delivery: 8.2 vs. 3.5
Karmon et al. (2009) [82]	Israel	CC	*n*=10,227	*n*=174,029	Stillbirth: 0.4 vs. 0.7; Macrosomia: 9.4 vs. 4.5; Preterm delivery: 8.3 vs. 7.7
Moses et al. (1995) [86]	Australia	ADIPS	*n*=138, Age: 29.5 (5.3)	*n*=144, Age: 28.2 (5.4)	Macrosomia: 8 vs. 17.4
Waters et al. (2016) [85]	North America	1) IADPSG 2) CC	1) *n*=878, Age: 31.0 (5.6), BMI: 31.5 (6.4) 2) *n*=261, Age: 32.3 (5.3), BMI: 31.6 (5.8)	*n*=5,020, Age: 30.1 (5.8), BMI: 28.2 (4.9)	1) Birth weight 90th percentile: 15.3 vs. 7.9 NICU admission: 8.1 vs. 6.3; Hyperbilirubinemia: 6.5 vs. 5; Preterm delivery: 7.7 vs. 6; Shoulder dystocia or birth injury: 3 vs. 1.8; Neonatal hypoglycemia: 2.9 vs. 1.3 2) Birth weight 90th percentile: 19.2 vs. 7.9; NICU admission: 9.6 vs. 6.3; Hyperbilirubinemia: 8.5 vs. 5; Preterm delivery: 13.8 vs. 6; Shoulder dystocia or birth injury: 2.3 vs. 1.8; Neonatal hypoglycemia: 3.1 vs. 1.3
Gu et al. (2019) [84]	China	WHO 1999	1) GDM with hypertensive disorders of pregnancy: *n*=91, Age: 33.8 (3.59), Pre-pregnancy BMI: 25.1 (3.64) 2) GDM without hypertensive disorders of pregnancy: *n*=1,172, Age: 33.3 (3.49), Pre-pregnancy BMI: 22.9 (3.24)	Non-GDM with hypertensive disorders of pregnancy: *n*=261, Age: 32.9 (2.68), Pre-pregnancy BMI: 22.2 (3.04) Non-GDM without hypertensive disorders of pregnancy: *n*=261, Age: 32.9 (2.84), Pre-pregnancy BMI: 21.4 (2.96)	1) Macrosomia: 28.6 vs. 7.4 2) Macrosomia: 19.1 vs. 11.2
Shand et al. (2008) [83]	Australia	ADIPS	*n*=16,727	*n*=349,933	Stillborn: 0.3 vs. 0.3; Birth trauma: 3 vs. 2.9; Neonatal hypoglycaemia: 19.1 vs. 1.6; Admission to NICU: 2.4 vs. 1.7; Birth weight percentile <10th centile: 8.5 vs. 9.8; Birth weight percentile >90th centile: 15.9 vs. 10.4
Anderberg et al. (2010) [80]	Sweden	WHO 1999	*n*=306, Age: 32 (18–46)	*n*=329, Age: 31 (20–42)	Born <37 gestational weeks: 8.9 vs. 2.7; LGA: 8.5 vs. 3.9; SGA: 2 vs. 1.5; Neonatal intensive care >1 day: 18.5 vs. 4.2
Avalos et al. (2013) [79]	Ireland	IADPSG	*n*=622, Age: 32.8	*n*=4,225, Age: 31 (4.9)	GDM without risk factor vs. GDM with risk factor vs. Non-GDM LGA: 16 vs. 24 vs. 15; NICU: 23 vs. 27 vs. 9
Wahabi et al. (2017) [78]	Saudi Arabia	WHO 2013	*n*=2,354, Age: 31.5 (5.9)	*n*=6,951, Age: 29.5 (5.7)	Macrosomia ≥4.0 kg: 4.8 vs. 2.5; Stillbirth: 0.9 vs. 0.9; Neonatal admission to NICU: 4.7 vs. 4.1
Meek et al. (2015) [77]	UK	1) IADPSG 2) NICE	1) *n*=387, Age: 32.6, BMI: 27.4 2) *n*=261, Age: 32.1, BMI: 25.5	*n*=2,406, Age: 31.4, BMI: 26	1) LGA: 29.7 vs.16.9; SGA: 6.4 vs. 9.6; Macrosomia: 28.9 vs. 16.8; Stillbirth: 0.3 vs. 0.2; Preterm: 7.5 vs. 5.3; Infant admission to NICU: 5.7 vs. 5.9 2) LGA: 11.5 vs.16.9; SGA: 12.6 vs. 9.6; Macrosomia: 9.2 vs. 16.8; Stillbirth: 0 vs. 0.2; Preterm: 8 vs. 5.3; Infant admission to NICU: 7.3 vs. 5.9
Boghossian et al. (2014) [76]	USA	ICD-9	*n*=1,279, Age:30.3 (4.9); Pre-pregnancy BMI: 28.9 (7.2)	*n*=58,224, Age: 28.1 (4.5), Pre-pregnancy BMI: 24.9 (5.6)	Preterm <37 wk: 11.4 vs. 7.3; SGA: 5.2 vs. 5.8; LGA: 17.1 vs. 9.2; Macrosomia: 11.8 vs. 7.4; Jaundice: 22.1 vs. 19; Shoulder dystocia: 3.1 vs. 2.2; Birth injury: 0.86 vs. 0.84; NICU admission: 12.4 vs. 7.6; Hypoglycemia: 1.6 vs. 1.9; RDS: 4.9 vs. 3.2; Stillbirth/neonatal mortality: 0.47 vs. 0.44

(*Continued to the next page*)

Supplementary Table 1. Continued

Study	Location	GDM diagnostic criteria	GDM characteristics	Non-GDM characteristics	Neonatal outcome in women with vs. without GDM, %
Kawakita et al. (2017) [75]	USA	ICD-9	*n*=11,327, Age: 30.8 (6.0), BMI: 34.1 (7.5)	*n*=208,355, Age: 27.4 (6.1), BMI: 30.6 (6.1)	RDS:4 vs. 3; Stillbirth: 0.4 vs. 0.4; NICU admission: 17.2 vs. 11.5
Brand et al. (2018) [74]	UK	Modified WHO 1999	White European: *n*=210, Age: 30.2 (5.4), BMI: 28.6 (6.3) South Asian: *n*=622, Age: 30.7 (5.3), BMI: 28.2 (5.8)	White European: *n*=4,537, Age: 26.6 (6.0), BMI: 26.5 (5.9) South Asian: *n*=5,336, Age: 27.7 (5.0), BMI: 25.2 (5.3)	White European SGA (<10th): 4.3 vs. 8.1; LGA (>90th):18.1 vs. 14.8 South Asian SGA (<10th): 11.4 vs. 16.7; LGA (>90th): 10.5 vs. 6.6
Kaul et al. (2015) [73]	Canada	CDA, 2013	GDM only: *n*=7,332, Age: 31.9 (5.5) GDM and overweight: *n*=1,399, Age: 31 (5.2)	*n*=213,765, Age: 28.6 (5.6)	GDM only vs. GDM and overweight vs. No GDM, not overweight LGA: 12.1 vs. 28.7 vs. 8.3; NICU stay: 16 vs. 20.2 vs. 10
Chen et al. (2019) [72]	Canada	ADA, 2003	First nations: *n*=1,828 Non-indigenous: *n*=1045,248	First nations: *n*=219,720 Non-indigenous: *n*=219,720	1) Preterm: 8.4 vs. 6.6; SGA: 1.9 vs. 4; LGA: 44 vs. 21.7 2) Preterm: 8.5 vs. 6.2; SGA: 7.7 vs. 8.7; LGA: 14.7 vs. 9
Feng et al. (2017) [71]	China	IADPSG	*n*=2,927	*n*=11,814	Macrosomia: 9.67 vs. 7.29; LGA: 9.33 vs. 5.87; Preterm birth: 6.29 vs. 4.98; SGA: 3.93 vs. 5.17
Kgosidialwa et al. (2015) [70]	Ireland	IADPSG	*n*=567, Age: 33.4 (4.9), BMI: 30.5 (6.1)	*n*=2,499, Age: 31.5 (5.2), BMI: 26.7 (4.8)	NICU: 15.5 vs. 7.6; LGA: 12.5 vs. 15.7; Macrosomia: 13.6 vs. 20.6; SGA: 4.1 vs. 5.2; Hypoglycemia: 3.2 vs. 0.3; Shoulder dystocia: 1.6 vs. 1.6; Prematurity: 5.5 vs. 3.4
Donovan et al. (2017) [69]	Canada	1) CDA 2) IADPSG	HAPO 1.75: *n*=4,308, Age: 31.2 (5.1) HAPO 2-1: *n*=5,528, Age: 31.6 (5.2) HAPO 2-2: *n*=3,252, Age: 32.1 (5.2)	Normal 50 g screen: *n*=144,191, Age: 28.8 (5.3) Normal 75 g OGTT: *n*=21,248, Age: 30.3 (5.3)	Normal 50 g screen LGA: 8.2; SGA: 9.7; Stillbirth: 0.2; Preterm: 6.3 Normal 75 g OGTT LGA: 10.5; SGA: 9.7; Stillbirth: 0.3; Preterm: 7.6 HAPO 1.75 LGA: 14.2; SGA: 8.3; Stillbirth: 0.3; Preterm: 9.4 HAPO 2-1 LGA: 11.8; SGA: 9.4; Stillbirth: 0.4; Preterm: 10 HAPO 2-2 LGA: 16.5; SGA: 8.5; Stillbirth: 0.2; Preterm: 12.6
Kieffer et al. (1999) [68]	Michigan	NDDG	*n*=19, Age: 29.4 (6.2), BMI: 28.7 (5.7)	*n*=353, Age: 24.79 (4.85), BMI: 25.1 (4.21)	SGA: 5.3 vs. 5.1; LGA: 21.1 vs.13
Ekeroma et al. (2015) [67]	New Zealand	1) NZSSD 2) IADPSG 3) ADIPS	1) *n*=381, Age: 31.7 (5.5), BMI: 31.8 (10.8) 2) *n*=238, Age: 31.4 (5.8), BMI: 32.9 (11.7) 3) *n*=608, Age: 31.5 (5.4), BMI: 30.5 (9.8)	*n*=1,672, Age: 30.0 (5.7), BMI: 30.7 (9.1)	1) Stillbirths: 0 vs. 1; Pre-term: 6 vs. 5; NICU admission: 8 vs. 6 2) Stillbirths: 0 vs. 1; Pre-term: 7 vs. 5; NICU admission: 6 vs. 6 3) Stillbirths: 0 vs. 1; Pre-term: 5 vs. 5; NICU admission: 8 vs. 6
Aung et al. (2015) [66]	Cook Islands	Modified IADPSG	*n*=94, Age: 36 (28–40), BMI: 34 (30–39)	*n*=28, Age: 24.79 (4.85), BMI: 31 (26–36)	Pre-term (n): 2 vs. 6; Admitted NICU (n): 8 vs. 22; Birth weight ≥4,000 g: 21 vs. 14
Erjavec et al. (2016) [65]	Croatia	1) WHO 1999 2) IADPSG	1) *n*=953, Age: 30.88 (5.23), BMI: 25.84 (5.28) 2) *n*=1,829, Age: 31.34 (5.19), BMI: 26.03 (5.64)	1) *n*=41,703, Age: 28.77 (5.23), BMI: 23.38 (3.99) 2) *n*=37,263, Age: 29.49 (5.33), BMI: 23.38 (4.11)	1) Birth weight >4,000 g: 17.2 vs. 11.9 2) Birth weight >4,000 g: 16.8 vs. 11.2

(*Continued to the next page*)

Supplementary Table 1. Continued

Study	Location	GDM diagnostic criteria	GDM characteristics	Non-GDM characteristics	Neonatal outcome in women with vs. without GDM, %
Gillespie et al. (2013) [64]	Ireland	WHO 2006	*n*=354, Age: 35.4 (6), BMI: 30.8 (7)	*n*=4,018, Age: 34.7 (5), BMI: 26.9 (5)	Neonatal unit admission: 28.7 vs. 10.1
Gortazar et al. (2019) [63]	Spain	NDDG	*n*=35,729, Age: 33.42	*n*=704,148, Age: 31.27	Macrosomia: 8.62 vs. 5.91; LGA: 18.33 vs. 12.71; SGA: 8.07 vs. 8.81; Preterm birth: 15.66 vs. 12.04
Zamstein et al. (2018) [62]	Israel	NDDG	GDM A1: *n*=9,460, Age: 32.1 (5.8) GDM A2: *n*=724, Age: 33.7 (5.6)	*n*=206,013, Age: 28 (5.7)	GDM A1 vs. GDM A2. vs. Normal SGA: 2.3 vs. 4.4 vs. 1.5; LGA: 11 vs. 18 vs. 3.8; Macrosomia: 10 vs. 13.3 vs. 4.4
Hedderson et al. (2003) [61]	California	1) NDDG 2) CC	1) *n*=1,523 2) *n*=840	*n*=38,515	1) SGA: 9.5 vs. 10.4; LGA: 15.4 vs. 9.2; Macrosomia: 16.8 vs. 13.9 2) SGA: 8.1 vs. 10.4; LGA: 19.4 vs. 9.2; Macrosomia: 23.4 vs. 13.9
Hosseini et al. (2018) [101]	Iran	IADPSG	Early-onset GDM: *n*=93, Age: 30.7 (4.6), Pre-pregnancy BMI: 26.5 (4.2) Late-onset GDM: *n*=78, Age: 31.1 (4.9), Pre-pregnancy BMI: 26.2 (4.7)	*n*=758, Age: 28.8 (4.6), Pre-pregnancy BMI: 24.2 (4.1)	Early-onset GDM vs. Late-onset GDM vs. Normal Macrosomia: 6.5 vs. 7.7 vs. 1.5; Preterm birth: 11.8 vs. 9 vs. 7.1; NICU admission: 12.9 vs. 3.8 vs. 4.4; Stillbirth: 0 vs. 0 vs. 0.3; Neonatal respiratory distress: 14 vs. 3.8 vs. 4.9
Hosseini et al. (2018) [101]	Iran	1) IADPSG 2) CC	1) *n*=78, Age: 18–45 2) *n*=35, Age: 18–45	1) *n*=758, Age: 18–45 1) *n*=801, Age: 18–45	1) Macrosomia (OR), 4.9; Preterm delivery (OR), 0.96 2) Macrosomia (OR), 13.3; Preterm delivery (OR), 0.65
Jain et al. (2016) [58]	India	DIPSI	*n*=8,000	*n*=7,641	Stillbirth: 3.2 vs. 1.3; LGA: 9 vs. 0.83; Jaundice: 5 vs. 1
Morikawa et al. (2017) [57]	Japan	IADPSG	*n*=13,037	*n*=223,108	Stillbirth: 46/68 vs. 1,562/2,001
Leybovitz-Haleluya et al. (2018) [56]	Israel	ACOG	GDM A1: *n*=9,460, Age: 32.1 (5.8) GDM A2: *n*=724, Age: 33.7 (5.6)	*n*=206,013, Age: 28 (5.7)	GDM A2 vs. GDM A1 vs. Normal SGA: 1.5 vs. 2.3 vs. 4.4; Macrosomia: 13.3 vs. 10 vs. 4.4; Preterm labor (<37 wk): 14.1 vs. 7.9 vs. 6.2; Non-reassuring fetal heart rate monitoring: 5.4 vs. 5.5 vs. 5
Jacobson et al. (1989) [55]	California	NDDG	*n*=97, Age: 28.8 (0.5), BMI: 27.6 (0.8)	*n*=2,107, Age: 26.3 (0.1), BMI: 22.8 (0.1)	LGA: 32 vs. 17.9; Hypoglycemia: 11.3 vs. 0.7; Hyperbilirubinemia: 10.3 vs. 1.8; Respiratory distress syndrome: 1 vs. 2.1; Stillborn: 0 vs. 1.8; Preterm labor: 6.2 vs. 7.1; Shoulder dystocia: 0 vs. 2
Schwartz et al. (1999) [54]	Washington	1) NDDG 2) CC	1) *n*=284 2) *n*=154	*n*=223,108	1) Stillbirth: 0.4 vs. 0.2; Birth weight ≥4,000: 18.2 vs. 16.5; Preterm birth: 6.3 vs. 4.4 2) Stillbirth: 0.6 vs. 0.2; Birth weight ≥4,000: 24.2 vs. 16.5; Preterm birth: 4.9 vs. 4.4
Pan et al. (2015) [53]	China	1) WHO 1999 2) IADPSG	1) *n*=257, Age: 29 (2.6), Pre-pregnancy BMI: 22.9 (3.5) 2) *n*=429, Age: 28.8 (2.9), Pre-pregnancy BMI: 23.9 (4)	*n*=16,173, Age: 28.4 (2.8), Pre-pregnancy BMI: 22.1 (3.3)	1) Stillbirth: 0.4 vs. 0.7; LGA: 11.6 vs. 9.5; Birth weight ≥40 kg: 8 vs. 8.6 2) Stillbirth: 0.2 vs. 0.7; LGA: 20.5 vs. 9.5; Birth weight ≥40 kg: 17.4 vs.8.6
Son et al. (2015) [52]	Korea	ICD-10	*n*=78,716, Age: 15–49	*n*=1,171,575, Age: 15–49	Preterm delivery: 2.96 vs. 2.04

(*Continued to the next page*)

Supplementary Table 1. Continued

Study	Location	GDM diagnostic criteria	GDM characteristics	Non-GDM characteristics	Neonatal outcome in women with vs. without GDM, %
von Katterfeld et al. (2012) [51]	Australia	ADIPS	Australian born *n*=4,765 CALD *n*=1,686 Non-CALD *n*=1,273	Australian born: *n*=142,537 CALD: *n*=23,541 Non-CALD: *n*=31,814	Australian born LGA: 14.6 vs. 10.3; Transfer to specialist care: 4 vs. 3.4; Preterm birth: 9.2 vs. 6.1; Shoulder dystocia: 2.5 vs. 1.7 CALD LGA: 13.2 vs. 10.6; Transfer to specialist care: 1.9 vs. 1.7; Preterm birth: 8.2 vs. 6.1; Shoulder dystocia: 2.6 vs. 1.7 Non-CALD LGA: 9.2 vs. 6.9; Transfer to specialist care: 4 vs. 2.9; Preterm birth: 8.3 vs. 6; Shoulder dystocia: 2.4 vs. 1.8
Savona-Ventura et al. (2013) [50]	Mediterranean countries	ADA 2003	*n*=112	*n*=1,139	Premature delivery <36 wk: 15.2 vs. 8.7; Shoulder dystocia: 0 vs. 0.3; Birth weight >4,000: 9.8 vs. 6.6; Respiratory distress syndrome: 7.1 vs. 4.8; Jaundice: 7.1 vs. 10.5; NICU admissions: 7.1 vs. 4; Prepartum death: 0 vs. 0.5
Sacks et al. (2015) [49]	California	IADPSG	1) GDM-1: *n*=771, Age: 30.9 (5.6) 2) GDM-2: *n*=1,121, Age: 31 (5.7)	*n*=7,943, Age: 26.3 (0.1)	GDM-1 vs. GDM-2 vs. Normal LGA: 15.6 vs. 20.2 vs. 9.9; Shoulder dystocia: 2.5 vs. 3.7 vs. 2.2; Respiratory distress: 1.8 vs. 2.5 vs. 1.6; Hyperbilirubinemia: 20.1 vs. 22.2 vs. 19.8; Neonatal hypoglycemia: 0.6 vs. 2.1 vs. 0.5; Preterm delivery: 6.9 vs. 10.8 vs. 6.2
Soliman et al. (2018) [48]	Qatar	IADPSG	*n*=3,027	*n*=8,995	Hypoglycemia: 3 vs. 0.6; Phototherapy: 8.9 vs. 7.2; Preterm: 9 vs. 6.4; Macrosomia: 6.8 vs. 5; NICU: 16 vs. 12.1; Stillborn: 0.3 vs. 0.8
Koivunen et al. (2017) [47]	Finland	Finland national criteria	*n*=6,679, Age: 31.02 (5.45), BMI before pregnancy: 28.2 (6.11)	*n*=52,386, Age: 29.39 (5.27), BMI before pregnancy: 23.8 (4.36)	LGA: 4.1 vs. 1.8; Admitted to neonatal ward: 12.9 vs. 7.9; Hypoglycemia: 22.1 vs. 2.7; Neonatal RDS: 0.2 vs. 0.3; Hyperbilirubinemia: 4.9 vs. 3.4; Fracture of the clavicle: 0.8 vs. 0.7; Erb's or Klumpke's palsy: 0.3 vs. 0.1 Diet 2010 vs. Insulin 2010 LGA: 3.7 vs. 7.1; Hypoglycemia: 20.4 vs.33.3; Neonatal RDS: 0.2 vs. 0.3; Transient tachypnea: 1.8 vs. 1.5; Hyperbilirubinemia: 4.4 vs. 7.7; Fracture of the clavicle: 0.9 vs. 0.3; Erb's or Klumpke's palsy: 0.3 vs. 0.3; Preterm births: 5.2 vs. 4.1
Xiong et al. (2001) [46]	Canada	CDA	*n*=2,755	*n*=108,664	Macrosomia: 9.3 vs. 5.9; SGA: 3.3 vs. 3.5; LGA: 10.3 vs. 3.1; Still birth: 0.3 vs. 0.3; Preterm birth: 10.4 vs. 7.5
Oster et al. (2014) [45]	Canada	CDA	*n*=1,224, Age: 28.8 (6.27)	*n*=26,793, Age: 24.7 (5.8)	High birth weight: 30 vs. 16.7; Stillbirth: 1.7 vs. 1.2; Neonatal intensive care unit admission: 16.1 vs. 8.6; Preterm: 16 vs. 9.2
Sugaya et al. (2000) [44]	Japan	1) JSOG 2) WHO 1998	1) *n*=55, Age: 29.7 (4.3), BMI: 26.2 (3.4) 2) *n*=51, Age: 32.8 (4.3), BMI: 26.5 (4.3)	*n*=281, Age: 30 (4.7), BMI: 25.5 (3.3)	1) HFD, birth weight >90: 14 vs. 7; LFD, birth weight <10:2 vs. 7; respiratory distress: 13 vs. 3; hyperbilirubinemia: 21 vs. 9 2) HFD, birth weight >90: 11 vs. 7; LFD, birth weight <10: 8 vs. 7; respiratory distress: 2 vs. 3; hyperbilirubinemia: 3 vs. 9
Fraser et al. (1994) [43]	Israel	NDDG	1) Jewish: *n*=346, Age: 27.9 (6.1) 2) Bedouin: *n*=96, Age: 33.4 (6.2)	1) Jewish: *n*=504, Age: 28.3 (5.4) 2) Bedouin: *n*=320, Age: 27.9 (6.1)	1) Jaundice: 11 vs. 6.2; Macrosomia: 9.2 vs. 4; Birth trauma: 3.8 vs. 1.2 2) Jaundice: 8.3 vs. 5; Macrosomia: 13.7 vs. 3.2; Birth trauma: 3.1 vs. 1.3
Kieffer et al. (2006) [42]	USA	ADA 2003	*n*=68, Age: 28.6 (0.6), BMI: 25.7 (0.2)	*n*=933, Age: 24.8 (0.2), BMI: 28.4 (0.8)	Still birth: 0 vs. 0.7; weighed 4,000 g: 13.2 vs. 8.8; Preterm birth: 19.1 vs. 5.9

(*Continued to the next page*)

Supplementary Table 1. Continued

Study	Location	GDM diagnostic criteria	GDM characteristics	Non-GDM characteristics	Neonatal outcome in women with vs. without GDM, %
Mahanta et al. (2014) [41]	India	Indian national criteria or modifies WHO 2006	*n*=28	*n*=749	Still birth: 0 vs. 0.7
Sletner et al. (2017) [40]	Norway	WHO 1999	Europe Mild: *n*=30, Age: 31.2 (29.5), BMI: 25.5 (23.8, 27.2) Moderate/severe: *n*=9, Age: 30.6 (27.6, 33.5), BMI: 30.5 (27.4, 33.6) South Asia Mild: *n*=9, Age: 30.7 (28.3, 33.0), BMI: 25.3 (23.2, 27.5) Moderate/severe: *n*=4,724, Age: 30.4 (28.0, 32.7), BMI: 22.7 (20.6, 24.9)	Europe: *n*=310, Age: 30.6 (30.1, 31.1), BMI: 24.3 (23.8, 24.8) South Asia: *n*=156, Age: 28.4 (27.7, 29.1), BMI: 23.7 (23.0, 24.3)	Europe Mild vs. Moderate/Severe vs. Non-GDM SGA: 13 vs. 0 vs. 8; LGA: 17 vs. 22 vs. 9; Preterm delivery: 13 vs. 11 vs. 4 South Asia Mild vs. Moderate/Severe vs. Non-GDM SGA: 0 vs. 21 vs. 26; LGA: 0 vs. 7 vs. 3; Preterm delivery: 15 vs. 0 vs. 7
Kong et al. (2019) [39]	Finland	ICD-9	*n*=98,568	*n*=542,735	LGA: 5.4 vs. 1.9; Premature offspring: 5.1 vs. 5.3
van Hoorn et al. (2002) [38]	Australia	ADIPS	*n*=51, Age: 30.9 (5.7), BMI: 31.5 (9.1)	*n*=258, Age: 24.9 (6.3), BMI: 25.5 (5.9)	Macrosomia: 7.8 vs. 7
Aberg et al. (2001) [37]	Sweden	Swedish national criteria	*n*=116, Age: 15–49	*n*=12,266, Age: 15–49	Deliveries at <37 wk: 8.4 vs. 5.7; birth weight of ≥4,500: 9.9 vs. 4.5; Stillborn (n): 2 vs. 0
Su et al. (2019) [36]	China	China national criteria	Underweight: *n*=1,466, BMI: 17.55 (0.79) Normal weight: *n*=6,905, BMI: 20.80 (1.21) Overweight: *n*=2,220, BMI: 23.86 (0.57) Obese: *n*=2,252, BMI: 27.21 (2.15)	Underweight: *n*=12,336, BMI: 17.54 (0.79) Normal weight: *n*=36,935, BMI: 20.54 (1.2) Overweight: *n*=6,654, BMI: 23.82 (0.56) Obese: *n*=4,730, BMI: 26.97 (1.97)	Preterm birth (aOR), 1.41 (1.29–1.55); LGA (aOR), 1.36 (1.29–1.44); SGA (aOR), 0.84 (0.75–0.95); Macrosamia (aOR), 1.59 (1.40–1.80)
Metcalfe et al. (2017) [35]	Canada	ICD-10	*n*=149,780	*n*=2,688,231	Preterm birth (<37 wk) (rate per 100 deliveries): 10.51 vs. 6.73
Carr et al. (2011) [34]	USA	ICD-9, 10	*n*=1,314, Age: 32.7 (5.7)	One abnormal: *n*=1,242, Age: 32.3 (5.3) Non abnormal: *n*=3,620, Age: 32 (5.7)	Preterm delivery: 18.6 vs.17.9 vs. 16.9

(*Continued to the next page*)

Supplementary Table 1. Continued

Study	Location	GDM diagnostic criteria	GDM characteristics	Non-GDM characteristics	Neonatal outcome in women with vs. without GDM, %
Lamminpaa et al. (2016) [33]	Finland	ICD-10	<35 yr: *n*=19,422 >35 yr: *n*=7,732	<35 yr: *n*=210,581 >35 yr: *n*=45,589	Normal glucose tol. vs. Diet-treated vs. Insulin-treated <35 yr Preterm delivery <28: 0.4 vs. 0.1 vs. 0.2; Preterm delivery 28–31: 0.6 vs. 0.4 vs. 0.3; Preterm delivery 32–36: 4.6 vs. 4.6 vs. 7.6; SGA (<5th percentile): 5.1 vs. 3.5 vs. 2.5; Admission to a neonatal unit: 10.7 vs. 14.6 vs. 28.8; Shoulder dystocia: 0.2 vs. 0.5 vs. 0.5; LGA (>95th percentile): 3.7 vs. 8.8 vs. 12.4 >35 yr Preterm delivery <28: 0.6 vs. 0.1 vs. 0.6; Preterm delivery 28–31: 0.8 vs. 0.4 vs. 0.5; Preterm delivery 32–36: 5.4 vs. 5.9 vs. 6.5; SGA (<5th percentile): 4.9 vs. 3.7 vs. 2.9; Admission to a neonatal unit: 12 vs. 15.7 vs. 2.5; Shoulder dystocia: 0.2 vs. 0.5 vs. 0.5; LGA (>95th percentile): 5 vs. 9.4 vs. 12.7
Black et al. (2010) [32]	California	IADPSG	1) Single isolated impaired glucose tolerance (i-IGT1): *n*=391, Age: 32.1 (5.4), Pregravid BMI: 28.1 (5.6) 2) Isolated impaired fasting glucose (i-IFG): *n*=886, Age: 30.4 (5.6), Pregravid BMI: 30.8 (7.1) 3) Double-isolated impaired glucose tolerance (i-IGT2): *n*=83, Age: 32.3 (5.2), Pregravid BMI: 27.5 (4.7) 4) IFG+IGT: *n*=331, Age: 32 (5.1), Pregravid BMI: 31.8 (7)	*n*=7,020, Age: 28.6 (5.9), Pregravid BMI: 26.9 (5.8)	1) LGA: 10 vs. 7.5; Preterm delivery: 12.8 vs. 6.6; Shoulder dystocia/birth injury: 4.6 vs. 3.8; Hyperbilirubinemia: 18.4 vs. 14 2) LGA: 16.7 vs. 7.5; Preterm delivery: 9.1 vs. 6.6; Shoulder dystocia/birth injury: 5.6 vs. 3.8; Hyperbilirubinemia: 14.4 vs. 14 3) LGA: 10.8 vs. 7.5; Preterm delivery: 18.1 vs. 6.6; Shoulder dystocia/birth injury: 6 vs. 3.8; Hyperbilirubinemia: 21.7 vs. 14 4) LGA: 20.5 vs. 7.5; Preterm delivery: 8.2 vs. 6.6; Shoulder dystocia/birth injury: 7 vs. 3.8; Hyperbilirubinemia: 13.6 vs. 14

GDM, gestational diabetes mellitus; IADPSG, International Association of the Diabetes and Pregnancy Study Groups; BMI, body mass index; NICU, neonatal intensive care unit; RDS, respiratory distress syndrome; LGA, large for gestational age; SGA, small for gestational age; CC, Carpenter-Coustan; ADIPS, The Australasian Diabetes in Pregnancy Society; WHO, World Health Organization; NICE, the National Institute for Clinical Excellence; ICD, International Classification of Diseases; CDA, Canadian Diabetes Association; ADA, American Diabetes Association; HAPO, Hyperglycemia and Adverse Pregnancy Outcomes; NDDG, National Diabetes Data Group; NZSSD, New Zealand Society for Study of Diabetes; OR, odds ratio; DIPSI, Diabetes in Pregnancy Study Group India; ACOG, American College of Obstetricians and Gynecologists; CALD, culturally and linguistically diverse; JSOG, Japan Society of Obstetrics and Gynecology; HFD, heavy for gestational date; LFD, light for gestational date; aOR, adjusted odds ratio.

Supplementary Table 2. Quality assessment of studies included using the Newcastle–Ottawa Quality Assessment Scale for cohort studies

Study	Selection				Comparability	Outcome			Total score
	Representativeness of the exposed cohort	Selection of the non-exposed cohort	Ascertainment of exposure	No outcome of interest at start of study	A: Study controls for age and/or BMI B: Study controls for other confounders	A: Doctor's diagnosis OR objective measurements B: Parent/self-reported doctor's diagnosis OR use of medication	Follow-up long enough for outcomes	Adequacy of follow-up of cohorts	
Capula et al. (2013) [81]	*	*	*	*	**	*	*	*	9*
Karmon et al. (2009) [82]	*	*	*	*		*	*	*	7*
Moses et al. (1995) [86]	*	*	*	*		*	*	*	7*
Waters et al. (2016) [85]	*	*	*	*	**	*	*	*	9*
Gu et al. (2019) [84]	*	*		*	**	*	*	*	8*
Anderberg et al. (2010) [80]	*	*		*	**	*	*	*	8*
Avalos et al. (2013) [79]	*	*	*	*		*	*	*	7*
Wahabi et al. (2017) [78]	*	*	*	*	**	*	*	*	9*
Meek et al. (2015) [77]	*	*	*	*	**	*	*	*	9*
Boghossian et al. (2014) [76]	*	*	*	*	**	*	*	*	9*
Kawakita et al. (2017) [75]	*	*	*	*		*	*	*	7*
Brand et al. (2018) [74]	*	*	*	*	**	*	*	*	9*
Kaul et al. (2015) [73]	*	*	*	*	**	*	*	*	9*
Chen et al. (2019) [72]	*	*	*	*	**	*	*	*	9*
Kgosidialwa et al. (2015) [70]	*	*	*	*		*	*	*	7*
Feng et al. (2017) [71]	*	*	*	*	*	*	*	*	8*
Gillespie et al. (2013) [64]	*	*	*	*	*	*	*	*	8*
Morikawa et al. (2017) [57]	*	*	*	*		*	*	*	7*
Schwartz et al. (1999) [54]	*	*	*	*		*	*	*	7*
Savona-Ventura et al. (2013) [50]	*	*	*	*		*	*	*	7*
Koivunen et al. (2017) [47]	*	*	*	*		*	*	*	7*
Donovan et al. (2017) [69]	*	*	*	*	*	*	*	*	8*
Kieffer et al. (1999) [68]	*	*	*	*	*	*	*	*	8*
Ekeroma et al. (2015) [67]	*	*	*	*	*	*	*	*	8*
Aung et al. (2015) [66]	*		*	*		*	*	*	6*
Gortazar et al. (2019) [63]	*	*	*	*	**	*	*	*	9*
Zamstein et al. (2018) [62]	*	*	*	*	**	*	*	*	9*
Hedderson et al. (2003) [61]	*	*	*	*	*	*	*	*	8*
Hosseini et al. (2018) [59]	*	*	*	*	*	*	*	*	8*

(*Continued to the next page*)

Supplementary Table 2. Continued

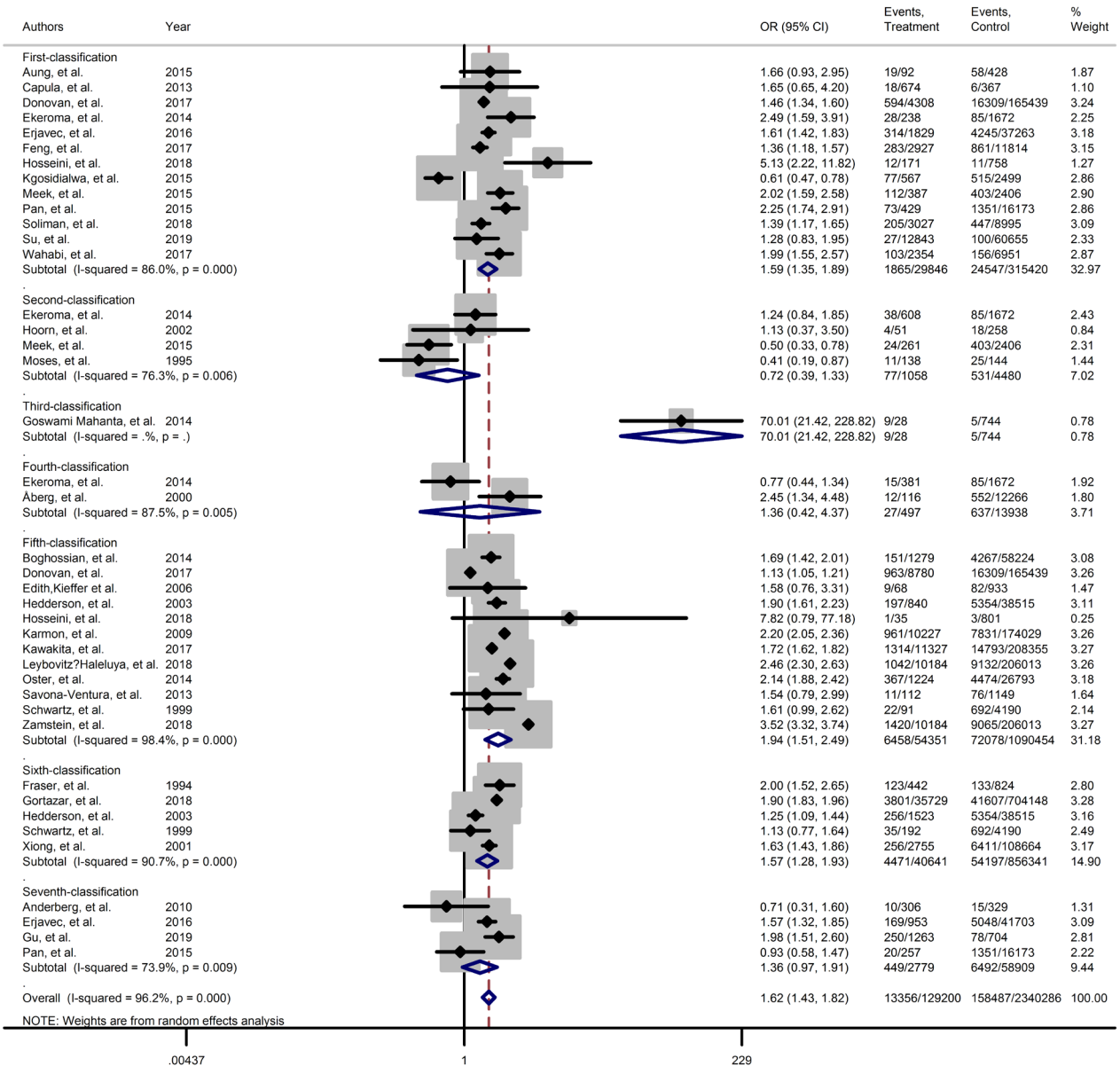
Study	Selection				Comparability	Outcome			Total score
	Representativeness of the exposed cohort	Selection of the non-exposed cohort	Ascertainment of exposure	No outcome of interest at start of study	A: Study controls for age and/or BMI B: Study controls for other confounders	A: Doctor's diagnosis OR objective measurements B: Parent/self-reported doctor's diagnosis OR use of medication	Follow-up long enough for outcomes	Adequacy of follow-up of cohorts	
Hosseini et al. (2018) [60]	*	*		*	**	*	*	*	9*
Jain et al. (2016) [58]	*	*		*		*	*	*	6*
Leybovitz-Haleluya et al. (2018) [56]	*	*	*	*	*	*	*	*	8*
Jacobson et al. (1989) [55]	*	*	*	*		*	*	*	7*
Pan et al. (2015) [53]	*	*	*	*	*	*	*	*	8*
Son et al. (2015) [52]	*	*	*	*	**	*	*	*	9*
von Katterfeld et al. (2012) [51]	*	*	*	*	**	*	*	*	9*
Sacks et al. (2015) [49]	*	*	*	*	**	*	*	*	9*
Oster et al. (2014) [45]	*	*		*	*	*	*	*	7*
Soliman et al. (2018) [48]	*	*		*		*	*	*	6*
Xiong et al. (2001) [46]	*	*	*	*	*	*	*	*	8*
Sugaya et al. (2000) [44]	*	*	*	*		*	*	*	7*
Fraser et al. (1994) [43]	*	*	*	*	*	*	*	*	8*
Kieffer et al. (2006) [42]	*	*	*	*	*	*	*	*	8*
Mahanta et al. (2014) [41]	*	*		*		*	*	*	6*
Kong et al. (2019) [39]	*	*	*	*	**	*	*	*	9*
Sletner et al. (2017) [40]	*	*	*	*	*	*	*	*	8*
Aberg et al. (2001) [37]	*	*	*	*		*	*	*	7*
van Hoorn et al. (2002) [38]	*	*	*	*		*	*	*	7*
Su et al. (2019) [36]	*	*	*	*	**	*	*	*	9*
Metcalfe et al. (2017) [35]	*	*	*	*	*	*	*	*	8*
Carr et al. (2011) [34]	*	*	*	*	*	*	*	*	8*
Lamminpaa et al. (2016) [33]	*	*	*	*	*	*	*	*	8*
Black et al. (2010) [32]	*	*	*	*	**	*	*	*	9*

It does not need to explain both * or ** . it means one score or two score.
BMI, body mass index.

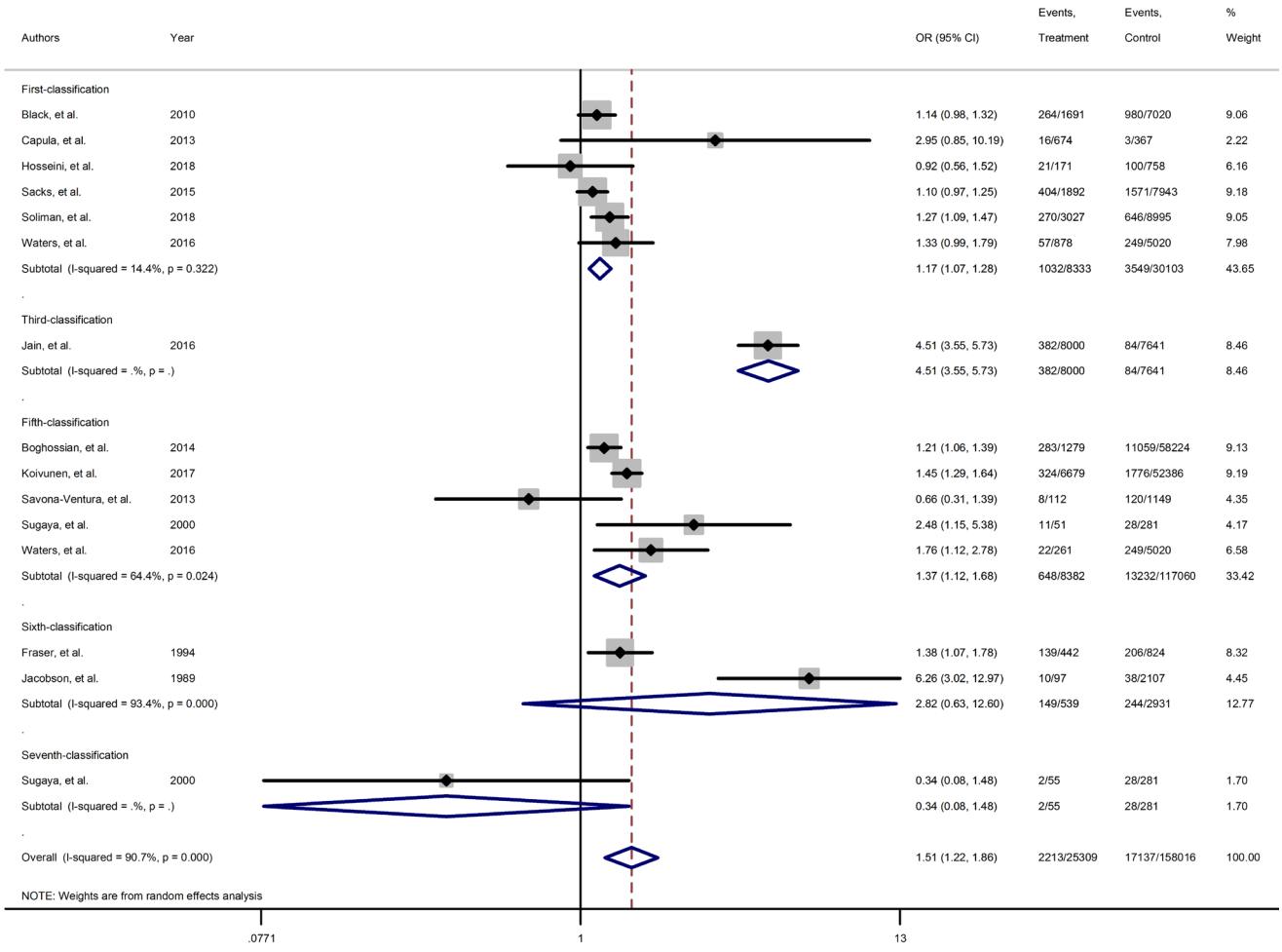
Supplementary Table 3. Quality assessment of included studies using the Newcastle–Ottawa Quality Assessment Scale for cross-sectional study

Study	Selection				Comparability	Outcome		Total scores
	Representativeness of the samples	Sample size	Non-responders	Ascertainment of the exposure (risk factor)	A: study controls for age and/or BMI B: control for any additional factor	Assessment of the outcome a) Independent blind assessment.** b) Record linkage.** c) Self report.*	Statistical test	
Erjavec et al. (2016) [65]	*		*	*		**	*	6*
Shand et al. (2008) [83]	*		*	*	**	**	*	8*

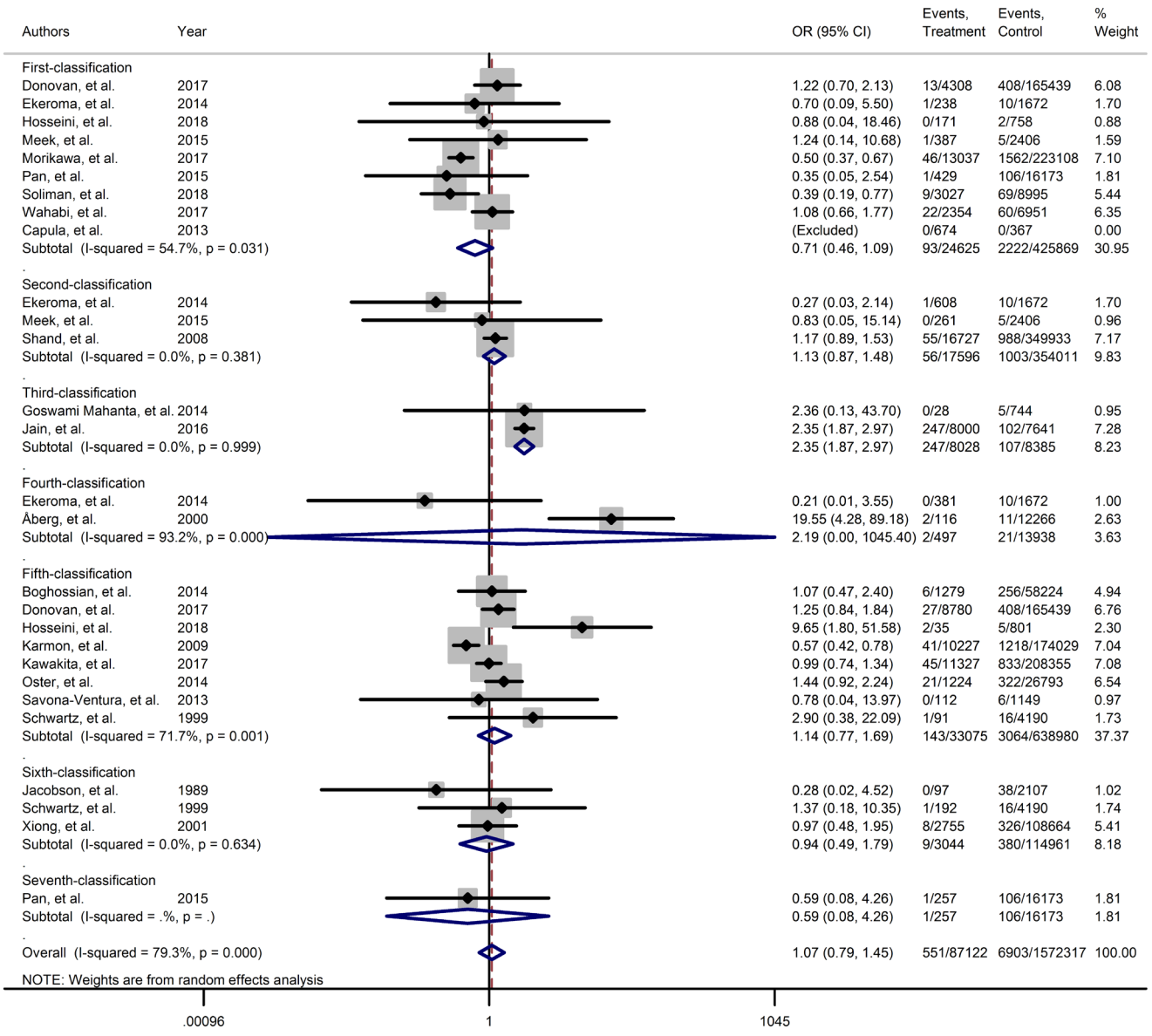
*: one score; **: two score.
BMI, body mass index.



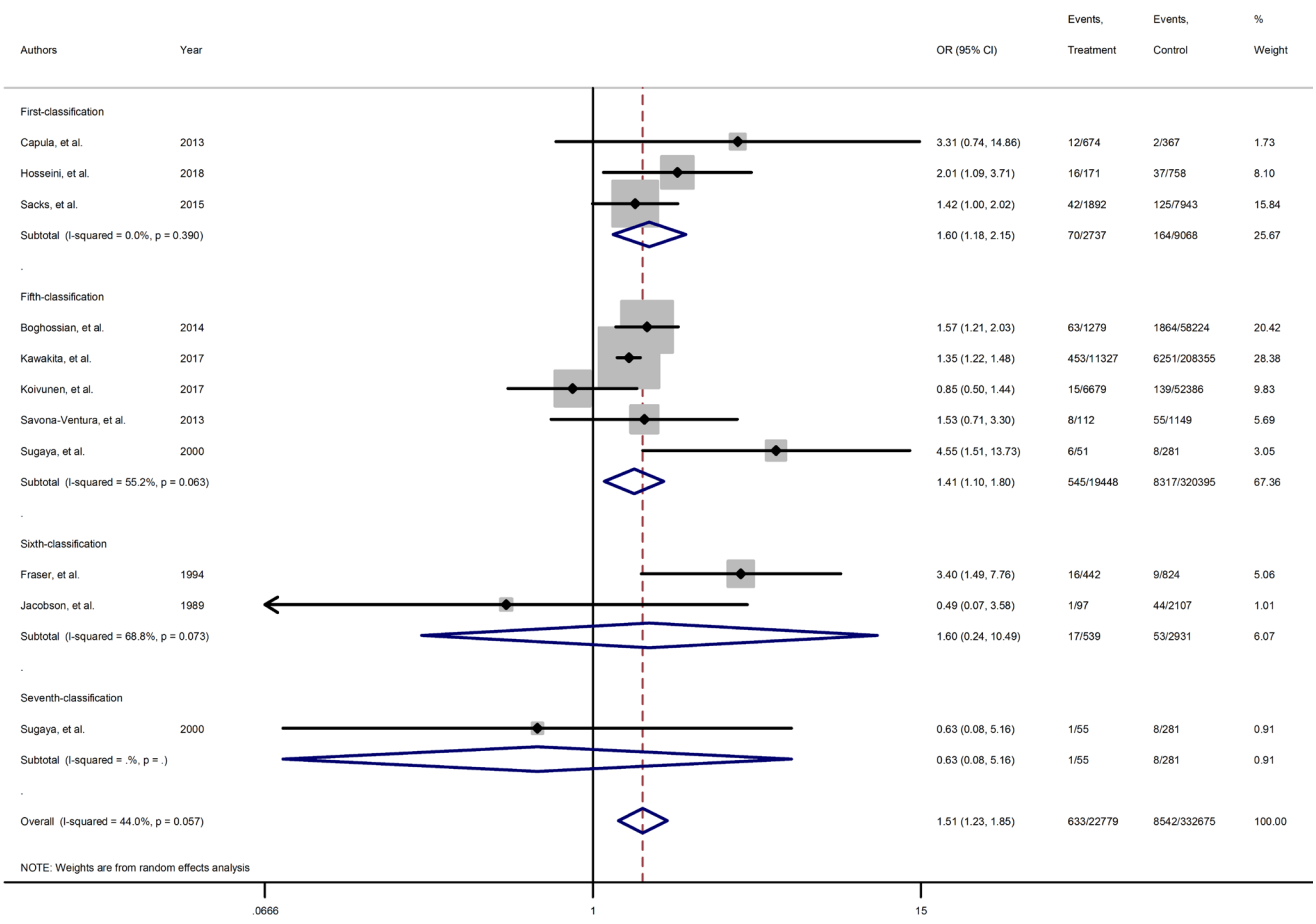
Supplementary Fig. 1. Forest plot of macrosomia obtained from Mantel–Haenszel method. Effect size (odds ratio [OR]) and 95% confidence intervals (CIs) for; pooled estimates of effect size are indicated by vertical points of diamonds and 95% CI are represented by horizontal points.



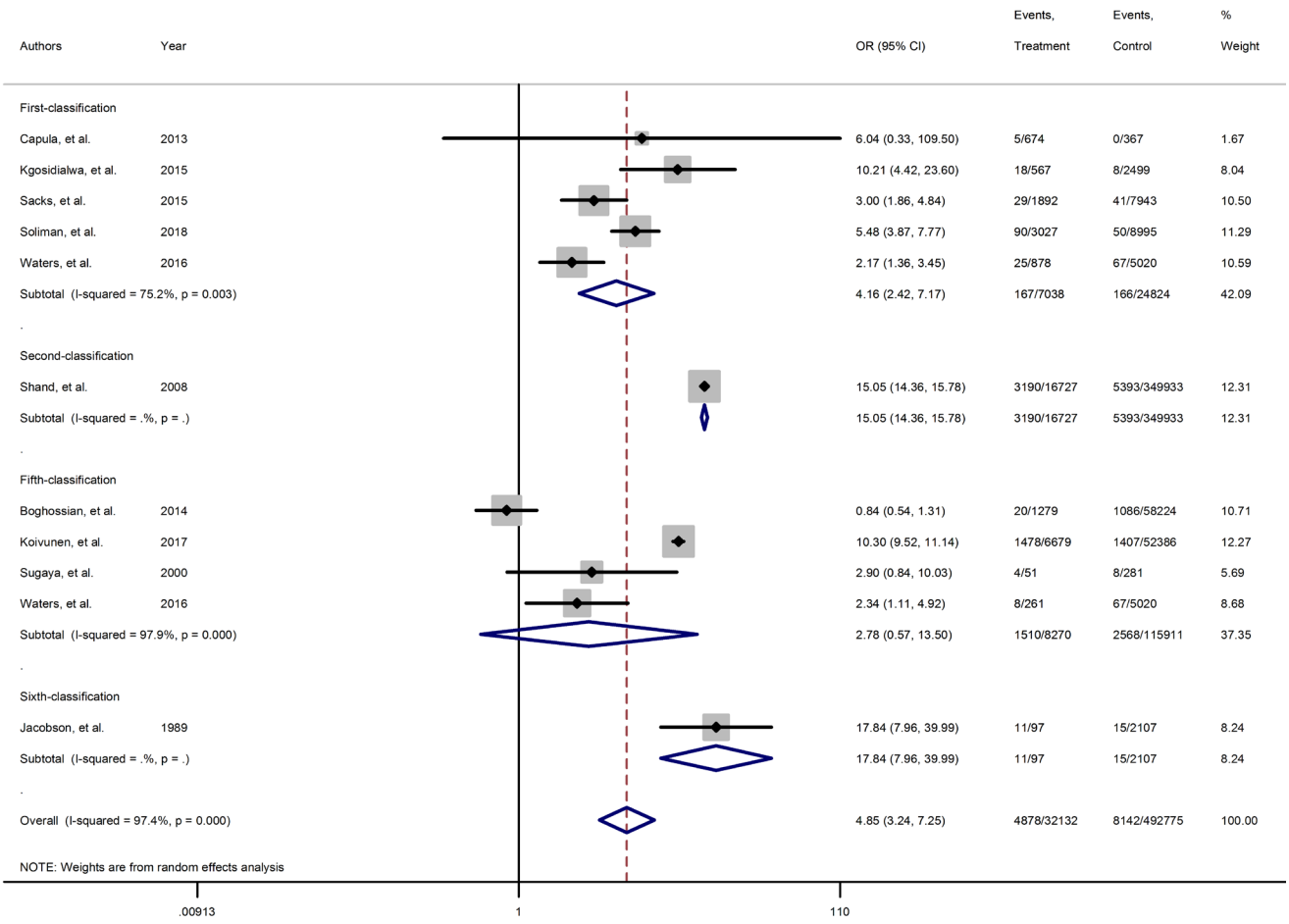
Supplementary Fig. 2. Forest plot of hyperbilirubinemia obtained from Mantel–Haenszel method. Effect size (odds ratio [OR]) and 95% confidence intervals (CIs) for; pooled estimates of effect size are indicated by vertical points of diamonds and 95% CI are represented by horizontal points.



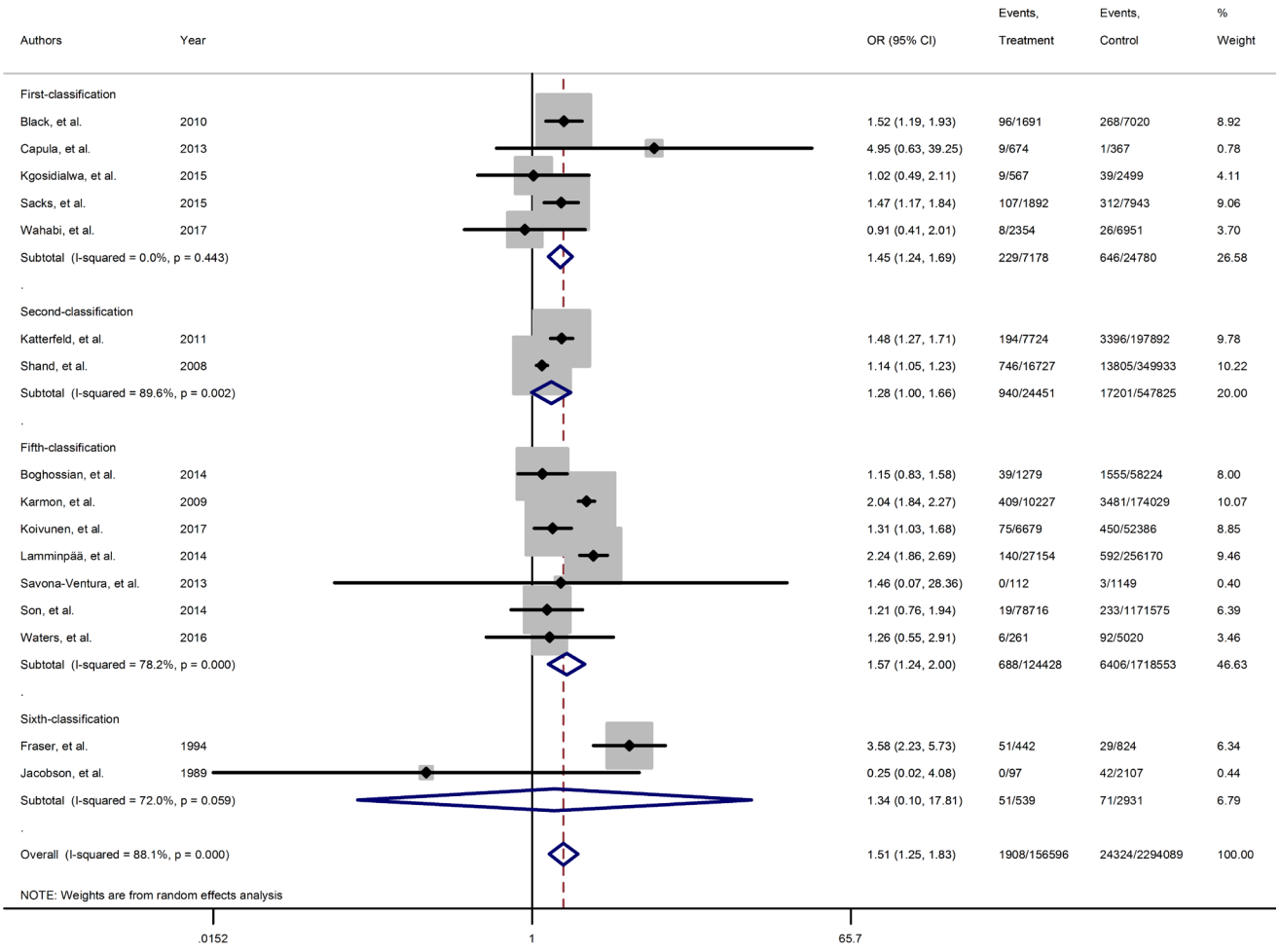
Supplementary Fig. 3. Forest plot of stillbirth obtained from Mantel–Haenszel method. Effect size (odds ratio [OR]) and 95% confidence intervals (CIs) for; pooled estimates of effect size are indicated by vertical points of diamonds and 95% CI are represented by horizontal points.



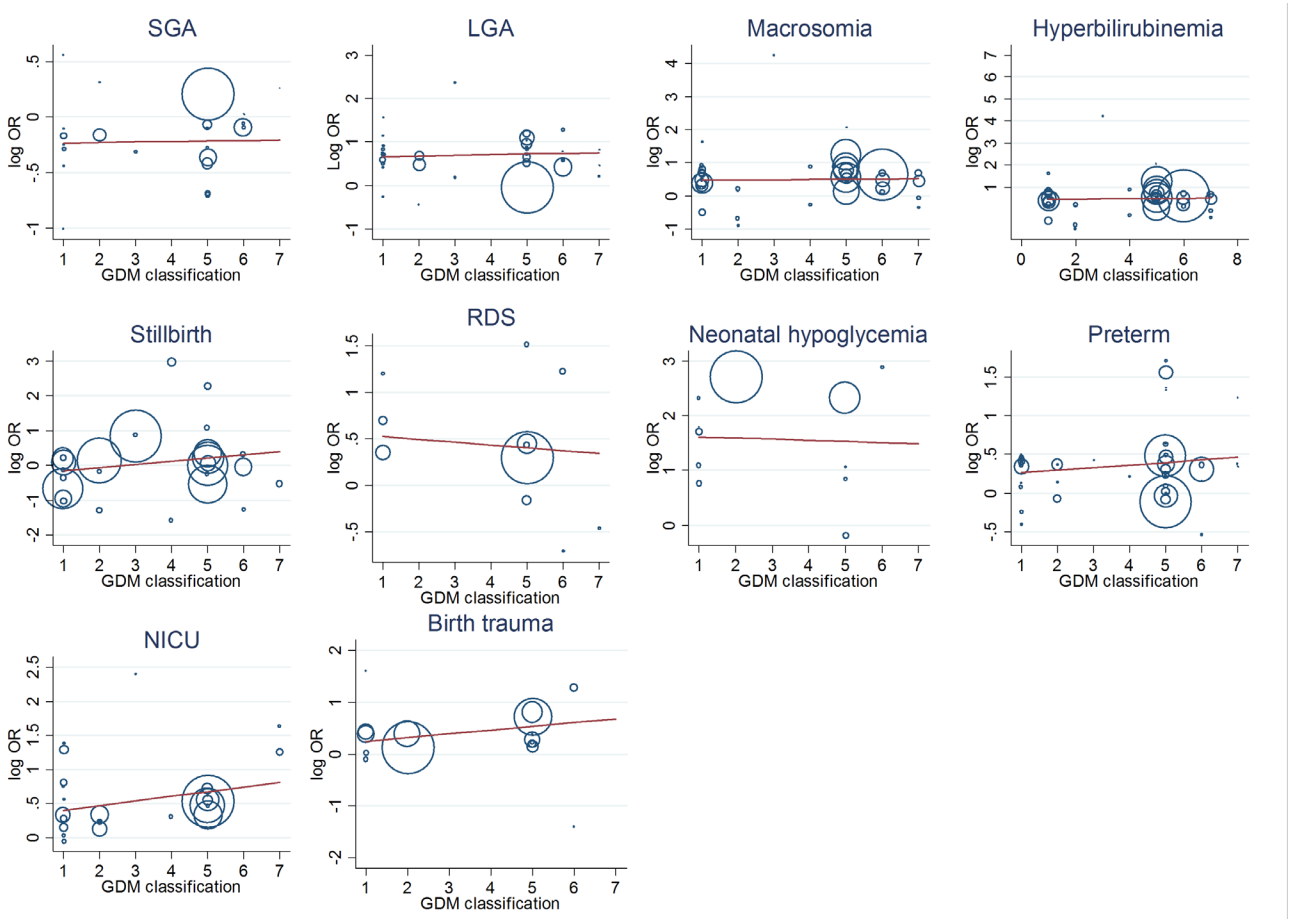
Supplementary Fig. 4. Forest plot of respiratory distress syndrome obtained from Mantel–Haenszel method. Effect size (odds ratio [OR]) and 95% confidence intervals (CIs) for; pooled estimates of effect size are indicated by vertical points of diamonds and 95% CI are represented by horizontal points.



Supplementary Fig. 5. Forest plot of neonatal hypoglycemia obtained from Mantel–Haenszel method. Effect size (odds ratio [OR]) and 95% confidence intervals (CIs) for; pooled estimates of effect size are indicated by vertical points of diamonds and 95% CI are represented by horizontal points.



Supplementary Fig. 6. Forest plot of birth trauma obtained from Mantel–Haenszel method. Effect size (odds ratio [OR]) and 95% confidence intervals (CIs) for; pooled estimates of effect size are indicated by vertical points of diamonds and 95% CI are represented by horizontal points.



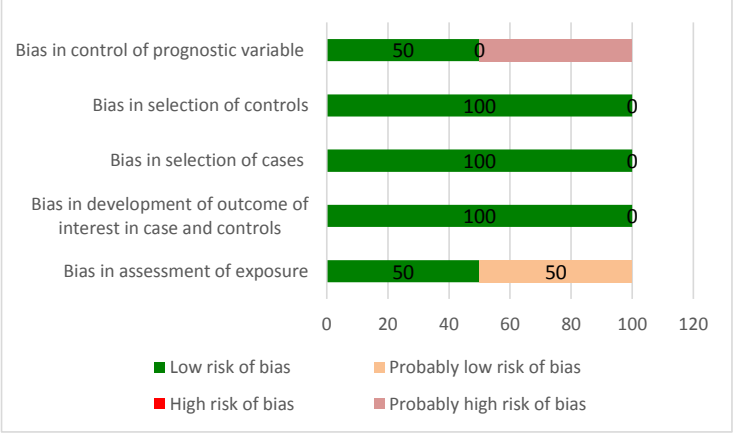
Supplementary Fig. 7. Bubble plots fitted meta-regression line for associations neonatal adverse outcomes of gestational diabetes mellitus (GDM) based on GDM diagnosis classification. OR, odds ratio; SGA, small for gestational age; LGA, large for gestational age; RDS, respiratory distress syndrome; NICU, neonatal intensive care unit.

Study	Bias in assessment of exposure (risk factor)	Bias in development of outcome of interest in case and controls	Bias in selection of cases	Bias in selection of controls	Bias in control of prognostic variable (without case and control matching or adjustment in statistical methods)
Erjavec et al. (2016) [65]	○	●	●	●	○
Shand et al. (2008) [83]	●	●	●	●	●

● Definitely no (low risk of bias) ○ Probably no

● Definitely yes (high risk of bias) ○ Probably yes

A



B

Supplementary Fig. 8. Risk of bias in each cross-sectional study (A) and overall (B).

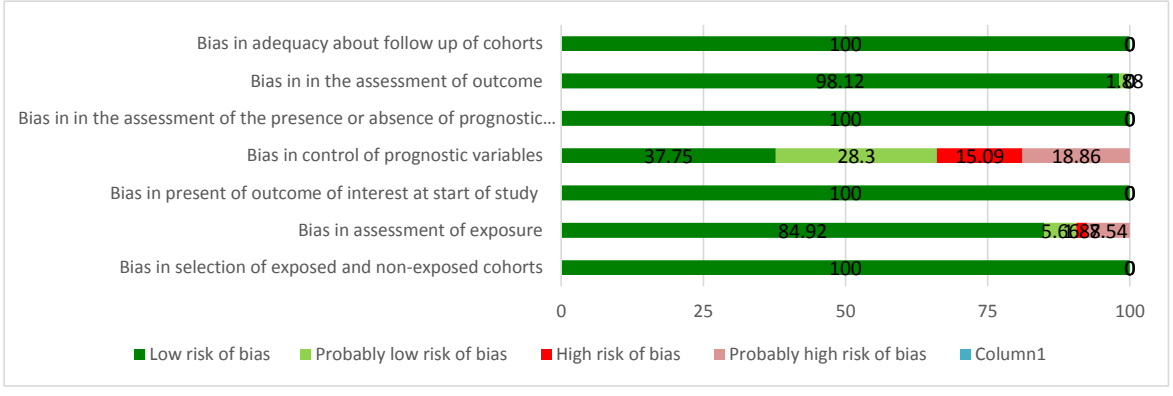
Study	Bias in selection of exposed and non-exposed cohorts	Bias in assessment of exposure	Bias in present of outcome of interest at start of study	Bias in control of prognostic variables (with matching or adjusting)	Bias in the assessment of the presence or absence of prognostic factors	Bias in the assessment of outcome	Bias in adequacy about follow-up of cohorts
Capula et al. (2013) [81]	+	+	+	+	+	+	+
Karmon et al. (2009) [82]	+	+	+	−	+	+	+
Moses et al. (1995) [86]	+	+	+	−	+	?	+
Waters et al. (2016) [85]	+	−	+	+	+	+	+
Gu et al. (2019) [84]	+	+	+	+	+	+	+
Anderberg et al. (2010) [80]	+	?	+	+	+	+	+
Avalos et al. (2013) [79]	+	+	+	×	+	+	+
Wahabi et al. (2017) [78]	+	+	+	+	+	+	+
Meek et al. (2015) [77]	+	+	+	+	+	+	+
Boghossian et al. (2014) [76]	+	+	+	+	+	+	+
Kawakita et al. (2017) [75]	+	?	+	+	+	+	+
Brand et al. (2018) [74]	+	+	+	+	+	+	+
Chen et al. (2019) [72]	+	+	+	+	+	+	+
Feng et al. (2017) [71]	+	+	+	?	+	+	+
Gillespie et al. (2013) [64]	+	+	+	−	+	+	+
Morikawa et al. (2017) [57]	+	+	+	×	+	+	+
Schwartz et al. (1999) [54]	+	+	+	×	+	+	+
Savona-Ventura et al. (2013) [50]	+	+	+	×	+	+	+
Koivunen et al. (2017) [47]	+	+	+	×	+	+	+
Kaul et al. (2015) [73]	+	+	+	+	+	+	+
Kgosidialwa et al. (2015) [70]	+	+	+	−	+	+	+
Donovan et al. (2017) [69]	+	+	+	?	+	+	+
Kieffer et al. (1999) [68]	+	+	+	?	+	+	+
Ekeroma et al. (2015) [67]	+	+	+	?	+	+	+
Aung et al. (2015) [66]	+	−	+	×	+	+	+
Gortazar et al. (2019) [63]	+	+	+	+	+	+	+
Zamstein et al. (2018) [62]	+	+	+	+	+	+	+
Hedderson et al. (2003) [61]	+	+	+	?	+	+	+
Hosseini et al. (2018) [59]	+	+	+	?	+	+	+
Hosseini et al. (2018) [60]	+	+	+	+	+	+	+
Jain et al. (2016) [58]	+	−	+	−	+	+	+
Leybovitz-Haleluya et al. (2018) [56]	+	+	+	?	+	+	+
Jacobson et al. (1989) [55]	+	+	+	−	+	+	+
Pan et al. (2015) [53]	+	+	+	?	+	+	+

(*Continued to the next page*)

Study	Bias in selection of exposed and non-exposed cohorts	Bias in assessment of exposure	Bias in present of outcome of interest at start of study	Bias in control of prognostic variables (with matching or adjusting)	Bias in the assessment of the presence or absence of prognostic factors	Bias in the assessment of outcome	Bias in adequacy about follow-up of cohorts
Son et al. (2015) [52]	●	●	●	●	●	●	●
von Katterfeld et al. (2012) [51]	●	●	●	●	●	●	●
Sacks et al. (2015) [49]	●	●	●	●	●	●	●
Soliman et al. (2018) [48]	●	●	●	●	●	●	●
Xiong et al. (2001) [46]	●	●	●	●	●	●	●
Oster et al. (2014) [45]	●	●	●	●	●	●	●
Sugaya et al. (2000) [44]	●	●	●	●	●	●	●
Fraser et al. (1994) [43]	●	●	●	●	●	●	●
Kieffer et al. (2006) [42]	●	●	●	●	●	●	●
Mahanta et al. (2014) [41]	●	●	●	●	●	●	●
Kong et al. (2019) [39]	●	●	●	●	●	●	●
Sletner et al. (2017) [40]	●	●	●	●	●	●	●
Aberg et al. (2001) [37]	●	●	●	●	●	●	●
van Hoorn et al. (2002) [38]	●	●	●	●	●	●	●
Su et al. (2019) [36]	●	●	●	●	●	●	●
Metcalfe et al. (2017) [35]	●	●	●	●	●	●	●
Carr et al. (2011) [34]	●	●	●	●	●	●	●
Lamminpaa et al. (2016) [33]	●	●	●	●	●	●	●
Black et al. (2010) [32]	●	●	●	●	●	●	●

● Definitely no (low risk of bias) ● Probably no
● Definitely yes (high risk of bias) ● Probably yes

A



B

Supplementary Fig. 9. Risk of bias in each included cohort study (A) and overall (B).