



Comparing IADPSG and NICE Diagnostic Criteria for GDM in Predicting Adverse Pregnancy Outcomes

Diabetes Care 2022;45:2046–2054 | <https://doi.org/10.2337/dc22-0579>

Yuanying He,¹ Ronald Ching Wan Ma,^{2,3,4}
H. David McIntyre,⁵ David A. Sacks,^{6,7}
Julia Lowe,⁸ Patrick M. Catalano,⁹ and
Wing Hung Tam¹

OBJECTIVE

To compare the performance of diagnostic criteria for gestational diabetes mellitus (GDM) proposed by the International Association of the Diabetes and Pregnancy Study Groups (IADPSG) with those endorsed by the National Institute for Health and Care Excellence (NICE) in predicting adverse pregnancy outcomes.

RESEARCH DESIGN AND METHODS

We performed a secondary data analysis of the Hyperglycemia and Adverse Pregnancy Outcomes (HAPO) study participants in five study centers. Logistic regression analyses were performed, and Akaike information criterion were applied for the comparison of different statistical prediction models. We further analyzed the performance by four racial/ethnic subgroups, namely, Whites, Hispanics, Asians, and Blacks.

RESULTS

Among all, IADPSG criteria diagnosed 267 (4.1%) more women with GDM, but predicted primary cesarean section (CS) and large for gestational age (LGA) and neonatal adiposity better than did NICE criteria after adjustment for potential confounders. Among Whites, IADPSG criteria diagnosed 65 (2.5%) more subjects with GDM and predicted LGA and neonatal adiposity better, but predicted hypertensive disorders, primary CS and clinical neonatal hypoglycemia worse. Among Hispanics, the IADPSG criteria diagnosed 203 (12.1%) more with GDM but performed better in predicting hypertensive disorders, LGA, neonatal adiposity, and hyperinsulinemia. Among Asians, the IADPSG criteria diagnosed 34 (2.0%) fewer subjects with GDM but predicted hypertensive disorders better in the unadjusted model. In Blacks, IADPSG criteria diagnosed 34 (10.5%) more women with GDM.

CONCLUSIONS

IADPSG criteria appear to be more favorable than NICE for identification of adverse pregnancy outcomes among Hispanic and Asian women, while they are comparable to NICE among White women.

The International Association of the Diabetes and Pregnancy Study Groups (IADPSG) diagnostic criteria for gestational diabetes mellitus (GDM) were adopted by the World Health Organization (WHO), the Australasian Diabetes in Pregnancy Society (ADIPS), and the International Federation of Gynecology and Obstetrics (FIGO) (1–3).

¹Department of Obstetrics and Gynaecology, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China

²Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China

³Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China

⁴Hong Kong Institute of Diabetes and Obesity, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China

⁵Mater Research, The University of Queensland, South Brisbane, Queensland, Australia

⁶Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, CA

⁷Department of Obstetrics and Gynecology, Keck School of Medicine, University of Southern California, Los Angeles, CA

⁸Faculty of Health and Medicine, University of Newcastle, Callaghan, New South Wales, Australia

⁹Department of Obstetrics and Gynecology, Mother Infant Research Institute, Tufts Medical Center, Boston, MA

Corresponding author: Wing Hung Tam, tamwh@cuhk.edu.hk

Received 24 March 2022 and accepted 14 June 2022

This article contains supplementary material online at <https://doi.org/10.2337/figshare.20199029>.

© 2022 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/journals/pages/license>.

However, they have not yet been universally adopted. In 2015, the National Institute for Health and Care Excellence (NICE) of the U.K. developed new diagnostic criteria for GDM based on a health economic modeling for immediate pregnancy complications. NICE (2015) proposed that a fasting plasma glucose (FPG) of 5.6 mmol/L (100 mg/dL) and a 2-h plasma glucose (PG) of 7.8 mmol/L (140 mg/dL) were more cost-effective compared with the IADPSG criteria based on the quality-adjusted life-year (QALY) and incremental cost-effectiveness ratio (ICER) (4). IADPSG criteria differ from those of NICE, with a substantially lower fasting glucose (5.1 mmol/L or 92 mg/dL) and a higher 2-h glucose value (8.5 mmol/L or 153 mg/dL), with an additional 1-h glucose level (10.0 mmol/L or 180 mg/dL).

Several studies have compared the performance of both criteria in predicting adverse pregnancy outcomes, but the conclusions are not uniform. The main limitations of previous studies were that the studies were mostly confined to Caucasian populations and that their treatment criteria varied.

In the original Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study, on which the IADPSG criteria were based (5), there was heterogeneity in the frequency of abnormal FPG, 1-h PG, and 2-h PG in women diagnosed with GDM across the 15 study centers (6). For example, an abnormal FPG level occurred in only 26% of women in the Hong Kong center but in >70% of women in Bellflower (7,8). Although FIGO advocated universal adoption of IADPSG criteria for the diagnosis of GDM, it remains uncertain whether the IADPSG criteria are superior to other existing diagnostic criteria in all ethnicities.

Because in clinical practice diagnosis of GDM is followed by treatment, it is extremely difficult to compare the application of different diagnostic criteria to the same patients who were treated for GDM by one set of criteria while they may or may not have been treated under another set of criteria. Furthermore, several studies have reported that women who met the IADPSG criteria but were diagnosed as non-GDM based on the NICE criteria had significantly increased risks of large for gestational age (LGA) infants and caesarean sections (CS) (9–12). Nevertheless, the diagnostic criteria in those studies

were either heterogenous or unknown. In the Cambridge historical cohort, the diagnostic criteria adopted for GDM management were WHO 1999 or modified WHO 1999, which were quite similar to the NICE 2015 criteria (10). A study of the Finnish population was, however, based on the local criteria (i.e., FPG 5.3 mmol/L, 1-h 10.0 mmol/L, 2-h 8.6 mmol/L) (11). Lastly, a study of the Indian population did not specify the criteria used for GDM treatment, so it is difficult to have a head-to-head comparison of the performance between the two criteria (12).

Our aim was to use a subset of the HAPO study data obtained from five study centers where all participants received no intervention for GDM and where both the clinicians and participants were blinded from the oral glucose tolerance test (OGTT) results.

RESEARCH DESIGN AND METHODS

Study Design

The original HAPO study protocol has been described previously elsewhere (13). We included data from five HAPO study centers: 1) Bellflower, CA; 2) Cleveland, OH; 3) Brisbane, Queensland, Australia; 4) Newcastle, New South Wales, Australia; and 5) Hong Kong, China.

All pregnant women booked for antenatal care in the study centers were eligible to participate unless they had one or more of the following exclusion criteria: age <18 years, a plan to undergo delivery at another hospital, an uncertain date of the last menstrual period and no ultrasonographic estimation between 6 and 24 weeks of gestational age, inability to complete the OGTT within 24–32 weeks of gestation, multiple pregnancy, conception by means of gonadotropin ovulation induction or in vitro fertilization, glucose testing before recruitment or a diagnosis of diabetes during the current pregnancy, diagnosis of diabetes before the current pregnancy and requiring treatment with medication, participation in another study that could interfere with the HAPO study, or infection with HIV or hepatitis B or C virus.

All participants underwent a 75-g OGTT at 24–32 weeks of gestation. Women with a FPG level of >105 mg/dL (5.8 mmol/L) or 2-h plasma glucose level of >200 mg/dL (11.1 mmol/L) were unblinded and received treatment for GDM accordingly. Otherwise, participants and health care

providers were blinded to the OGTT results, and participants received no treatment. All maternal and umbilical cord blood samples were analyzed at the central laboratory of the HAPO study.

Self-reported race/ethnicity (i.e., White, Hispanic, Asian [majority Chinese], Black, and others/miscellaneous), maternal age, gestational age at delivery, BMI at OGTT, smoking status, and parity (nulliparity vs. multiparity) were extracted as demographic and clinical characteristics. Neonatal measurements were obtained by a standardized method within 72 h after the delivery. Infants' percentage of body fat was estimated as $(0.39055 [\text{birth weight}] + 0.0453 [\text{flank skinfold}] - 0.03237 [\text{length}] + 0.54657)/[\text{birth weight}]$, as previously described (14).

Diagnostic Definition of GDM by NICE and IADPSG Criteria

Women whose FPG was ≥ 5.6 mmol/L (100 mg/dL) and/or 2-h PG was ≥ 7.8 mmol/L (140 mg/dL) were defined as GDM by NICE criteria, while those whose FPG was ≥ 5.1 mmol/L (92 mg/dL) and/or 1-h PG was ≥ 10.0 mmol/L (180 mg/dL) and/or 2-h PG was ≥ 8.5 mmol/L (153 mg/dL) were defined as GDM by IADPSG criteria. Furthermore, all women were classified under one of the four categories: 1) non-GDM by both IADPSG and NICE; 2) GDM by IADPSG but non-GDM by NICE; 3) GDM by NICE but non-GDM by IADPSG; and 4) GDM by both the IADPSG and NICE criteria.

Outcome Variables

We compared participants who remained blinded on six adverse pregnancy outcomes: 1) hypertensive disorder of pregnancy, namely, gestational hypertension (GH) or preeclampsia (PE); 2) primary CS; 3) LGA infants (infants born with birth weight >90th percentile adjusted for ethnicity, study center, gestational age, parity, and infant's sex); 4) neonatal adiposity (percentage body fat >90th percentile adjusted for ethnicity, study center, gestational age, parity, and infant's sex); 5) clinical neonatal hypoglycemia (being present if there was a notation of neonatal hypoglycemia in the medical record and there were symptoms or treatment with a glucose infusion or a local laboratory report of a glucose value of ≤ 1.7 mmol/L in the first 24 h after birth or ≤ 2.5 mmol/L after the first 24 h [13]); and 6) neonatal hyperinsulinemia (umbilical cord blood

C-peptide level >90th percentile; i.e., 1.7 $\mu\text{g/L}$ (15,16), among all blinded participants and among four ethnic subgroups (i.e., Whites, Hispanics, Asians, and Blacks).

Statistical Analyses

Continuous variables are expressed as mean \pm SD and were compared by using the Kruskal-Wallis test. Post hoc pairwise analyses were corrected by the Wilcoxon test when the null hypotheses were rejected. Categorical variables are expressed as n (%) and were compared by the McNemar test, χ^2 test, or Fisher exact test, as appropriate, for between-group comparisons and specified in legends of the tables. If the null hypothesis were rejected, post hoc pairwise comparisons were further conducted, corrected by the false discovery rate. Logistic regressions were used to explore the association of GDM by IADPSG or NICE criteria with adverse pregnancy outcomes. Data were adjusted for maternal age, parity, race/ethnicity, and study center (model 1), and model 1 plus maternal height, BMI and gestational age at OGTT, smoking status, alcohol use, maternal urinary tract infection (for GH/PE only), hospitalization before delivery, and mean arterial pressure at OGTT (except for GH/PE), family history of diabetes, and baby's sex (model 2) (13). We also applied Akaike information criterion (AIC) for the comparison of different statistical prediction models in the logistic regression analysis (17). A model with a lower AIC value of >4 was considered as significantly better in performance (17). P values <0.05 were used to indicate significance for two-tailed statistical test results. All statistical analyses were performed by using R 4.0.3 software (downloadable at www.r-project.org).

RESULTS

Among 6,544 women recruited from the five HAPO study centers, 1,223 (18.7%) and 956 women (14.6%) were diagnosed with GDM by the IADPSG and NICE criteria, respectively (Table 1). Among the 1,732 participants in the Hong Kong center, 1,595 were Chinese, suggesting at least 92% of the Asians in this cohort were of Chinese race/ethnicity. IADPSG diagnosed significantly more women with GDM than did NICE criteria among

Whites (15.7% vs. 13.2%, $P < 0.001$), Hispanics (26.4% vs. 14.3%, $P < 0.001$), and Blacks (20.7% vs. 10.2%, $P < 0.001$), but less GDM among Asians (15.7% vs. 17.7%, $P = 0.011$). When comparing by study centers, IADPSG diagnosed more women with GDM than did NICE among women from the two U.S. centers (i.e., Bellflower and Cleveland), but less GDM among women from the Hong Kong center. (Table 1). There were no differences in the incidence of GDM by the two diagnostic criteria in either of the Australian centers. In this multiracial and multiethnic cohort, Hispanic women have the highest GDM rate if IADPSG criteria were adopted, while Asian women (majority Chinese) have the highest rate if NICE diagnostic criteria were used (Supplementary Fig. 1); whereas, the frequency of women being non-GDM by both criteria was lowest in Hispanics but highest in Whites. In contrast, the frequency of women with GDM by both criteria was highest in Hispanics but lowest in Whites (Supplementary Table 1).

Figure 1 shows the frequency distribution of FPG and 1-h and 2-h PG levels in Whites, Hispanics, Asians, and Blacks. Asians, the majority of whom were Chinese, had a lower rate (2.5%) of FPG level ≥ 5.1 but <5.6 mmol/L (i.e., above the threshold for IADPSG but below the threshold of NICE) than Whites (5.3%, $P < 0.001$), Hispanics (10.2%, $P < 0.001$), and Blacks (7.7%, $P < 0.001$) (Fig. 1A, D, G, and J). In contrast, Hispanics had a higher rate (13.1%) of 1-h PG ≥ 10.0 mmol/L (i.e., above the threshold for

IADPSG) than Whites (7.4%, $P < 0.001$), Asians (9.8%, $P = 0.010$), and Blacks (7.7%, $P = 0.014$) (Fig. 1B, E, H, and K). Moreover, Hispanics had a significantly lower rate (4.9%) of 2-h PG ≥ 7.8 but <8.5 mmol/L (i.e., above the threshold for NICE but below the threshold of IADPSG) than Whites (6.5%, $P = 0.045$) and Asians (8.1%, $P < 0.001$) (Fig. 1C, F, and I). Similarly, fewer Blacks (3.1%) had 2-h PG ≥ 7.8 but <8.5 mmol/L than Whites ($P = 0.041$) and Asians ($P = 0.006$) (Fig. 1C, I, and L). On the other hand, Whites had a lower rate of FPG level ≥ 5.1 but <5.6 mmol/L than Hispanics ($P < 0.001$) (Fig. 1A and D) and a lower rate of 1-h PG ≥ 10.0 mmol/L than Asians ($P = 0.013$) (Fig. 1B and H).

After the exclusion of 147 women whose data were unblinded, data from 6,397 women were used to analyze the relationship between the diagnosis of GDM by each of the criteria and adverse pregnancy outcomes. Their maternal and neonatal characteristics compared by races/ethnicities are summarized in Supplementary Table 2. Hispanics had the lowest frequency of nulliparous women, while Asians had the highest frequency. Moreover, Asian women were the oldest and had the lowest BMI at OGTT, while Black women were the youngest and had the highest BMI at OGTT (Supplementary Table 2). Incidences of adverse pregnancy outcomes in women diagnosed with GDM by IADPSG and NICE criteria, respectively, in overall and in each race/ethnicity are tabulated in Supplementary Table 3.

Table 1—The frequencies of GDM diagnosed according to IADPSG and NICE criteria by races/ethnicities and HAPO study centers

	Participants	IADPSG	NICE	P
Overall	6,544 (100)	1,223 (18.7)	956 (14.6)	<0.001
Race/ethnicities				
White	2,583 (39.5)	406 (15.7)	341 (13.2)	<0.001
Hispanic	1,677 (25.6)	442 (26.4)	239 (14.3)	<0.001
Asian*	1,732 (26.5)	272 (15.7)	306 (17.7)	0.011
Black	324 (5.0)	67 (20.7)	33 (10.2)	<0.001
Other	228 (3.5)	36 (15.8)	37 (16.2)	0.835
Centers				
Bellflower	1,981 (30.3)	505 (25.5)	271 (13.7)	<0.001
Cleveland	797 (12.2)	199 (25.0)	134 (16.8)	<0.001
Brisbane	1,444 (22.1)	179 (12.4)	171 (11.8)	0.505
Newcastle	668 (10.2)	102 (15.3)	93 (13.9)	0.292
Hong Kong	1,654 (25.3)	238 (14.4)	287 (17.4)	<0.001

Data are expressed as n (%). Frequencies were compared by the McNemar test. *More than 92% of the Asians in the cohort were Chinese ethnicity.

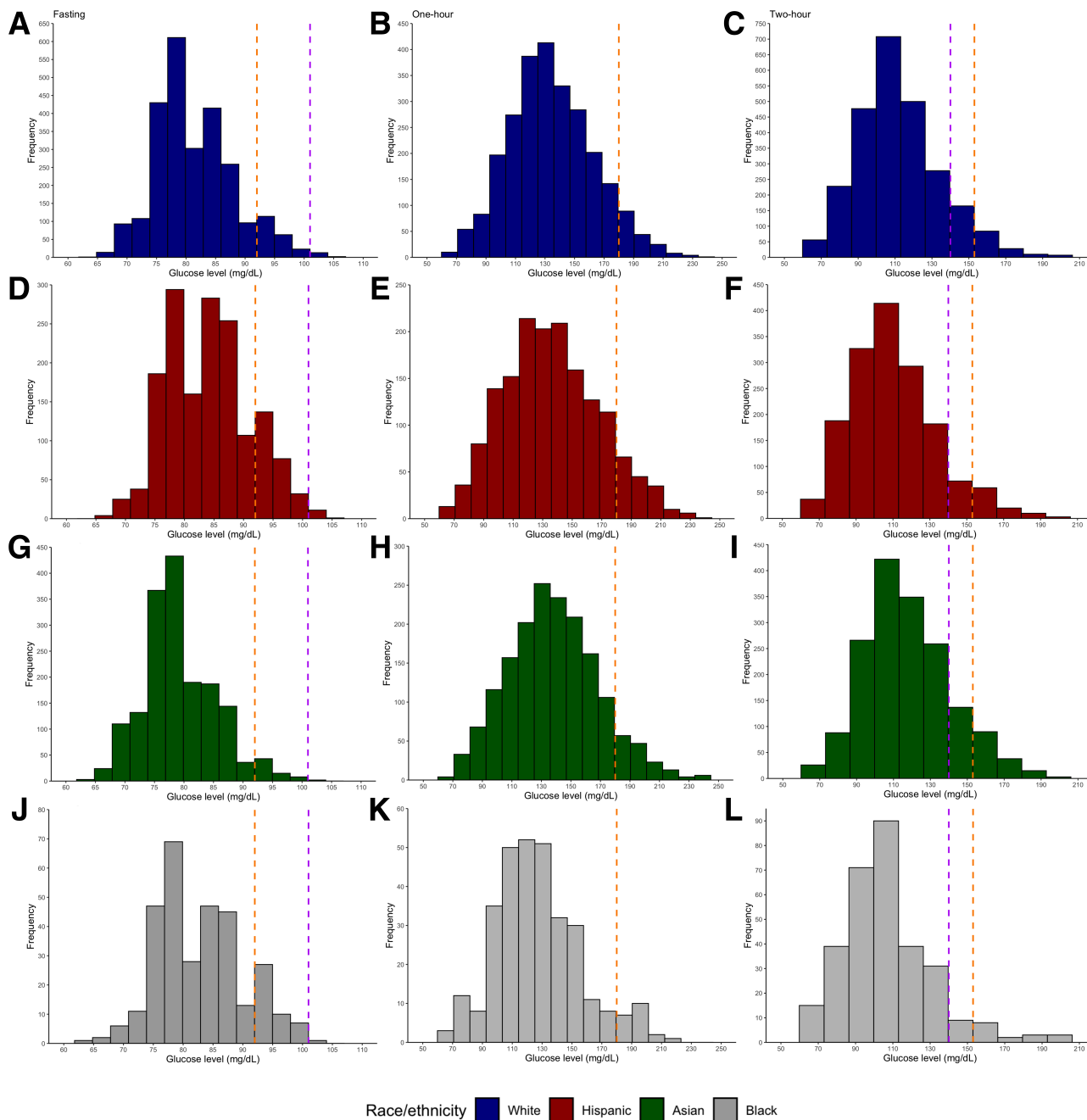


Figure 1—The frequency distribution of plasma glucose levels at OGTT among the White (A, B, and C), Hispanic (D, E, and F), Asian (>92% of the Asians in the cohort were Chinese ethnicity) (G, H, and I), and Black (J, K, and L) women in the HAPO study subcohort. The dashed lines in orange and purple indicate the thresholds of the IADPSG and NICE criteria, respectively.

Supplementary Table 4 compares the performance of IADPSG with that of NICE criteria in the prediction of maternal and neonatal adverse outcomes among all women whose OGTT result remained blinded. The IADPSG performed better than NICE criteria in predicting hypertensive disorders, primary CS, LGA infants, higher neonatal adiposity, and hyperinsulinemia, indicated by AIC, after adjustment for maternal age, parity, race/ethnicity, and study center (model 1).

However, the performance of both IADPSG and NICE were attenuated with further adjustment for confounders in model 2, which weakened the better performance of IADPSG in predicting hypertensive disorders and neonatal hyperinsulinemia (model 2). Clinical neonatal hypoglycemia was only predicted by the IADPSG, but not the NICE criteria in the unadjusted regression model. However, the performances of the IADPSG and NICE criteria were attenuated and enhanced,

respectively, after adjustment for the potential confounders (model 1 and 2), which in contrast, made the clinical neonatal hypoglycemia only predicted by the NICE but not the IADPSG criteria (Supplementary Table 4).

Table 2 shows the result of comparisons in Whites. IADPSG criteria performed better than NICE criteria in predicting LGA infants, indicated by AIC, after adjustment for confounders in model 2. On the other hand, neonatal adiposity was predicted

Table 2—Associations of GDM diagnosed by the IADPSG and NICE criteria with adverse pregnancy outcomes in Whites, whose OGTT remained blinded (n = 2,551)

	Adjusted											
	Unadjusted				Model 1				Model 2			
	IADPSG OR (95% CI)	AIC	NICE OR (95% CI)	AIC	IADPSG OR (95% CI)	AIC	NICE OR (95% CI)	AIC	IADPSG OR (95% CI)	AIC	NICE OR (95% CI)	AIC
Maternal adverse outcomes												
GH/PE	1.99 (1.52–2.62)§	2,109.7	1.85 (1.38–2.49)§	2,117.0	1.81 (1.37–2.40)§	2,021.6	1.79 (1.32–2.43)§	2,024.6	1.28 (0.95–1.73)	1,896.5	1.47 (1.06–2.03)*	1,893.9
Primary CS	1.61 (1.23–2.11)§	2,204.9	1.75 (1.31–2.33)§	2,202.2	1.68 (1.26–2.24)§	2,046.8	1.64 (1.21–2.21)‡	2,049.1	1.30 (0.96–1.76)	2,011.0	1.37 (1.00–1.87)*	2,010.1
Neonatal outcomes												
LGA	1.99 (1.44–2.71)§	1,595.0	1.42 (0.97–2.02)	1,608.1	2.11 (1.52–2.90)§	1,596.2	1.51 (1.03–2.16)*	1,610.6	2.02 (1.42–2.83)§	1,535.9	1.64 (1.10–2.39)*	1,544.9
Adiposity	1.92 (1.32–2.74)§	1,247.3	1.30 (0.84–1.97)	1,257.2	2.06 (1.41–2.96)§	1,247.7	1.40 (0.90–2.13)	1,258.7	1.80 (1.21–2.63)‡	1,244.4	1.33 (0.84–2.04)	1,251.1
Clinical hypoglycemia	1.76 (0.69–3.93)	328.5	3.12 (1.35–6.68)‡	323.3	1.96 (0.76–4.43)	322.5	3.42 (1.45–7.46)‡	317.1	1.61 (0.60–3.83)	330.0	2.89 (1.19–6.54)*	325.6
Hyperinsulinemia	2.39 (1.67–3.39)§	1,207.4	2.53 (1.73–3.64)§	1,207.2	2.31 (1.59–3.32)§	1,170.2	2.52 (1.70–3.68)§	1,169.0	2.00 (1.36–2.91)§	1,176.5	2.29 (1.53–3.39)§	1,173.1

Model 1 was adjusted for maternal age, parity, ethnicity, and study center. Model 2 was adjusted for covariates in model 1 plus maternal height, BMI, and gestational age at OGTT, smoking status, alcohol use, maternal urinary tract infection (for GH/PE only), hospitalization before delivery, and mean arterial pressure at OGTT (except for GH/PE), family history of diabetes, and baby's sex. OR, odds ratio. *P < 0.05; †P < 0.005; ‡P < 0.001.

only by the IADPSG, whereas hypertensive disorders, primary CS, and clinical neonatal hypoglycemia were predicted only by the NICE criteria, based on the P value (Table 2).

In Hispanics, GDM by IADPSG criteria predicted all individual outcomes, except primary CS and neonatal hypoglycemia, while NICE criteria only identified women with LGA infants, after adjustment for confounders in model 2. Moreover, the IADPSG criteria also performed better in predicting LGA infants indicated by the AIC. Meanwhile, primary CS could only be predicted by IADPSG, but not NICE criteria, and confined to unadjusted model only (Table 3).

In Asians, the majority of whom were Chinese, hypertensive disorders were predicted only by the IADPSG in an unadjusted model. Nevertheless, the better performance of the IADPSG was attenuated after adjustment for the potential confounders (model 1 and 2). On the other hand, NICE performed better in predicting neonatal adiposity, indicated by the AIC value, after adjustment for confounders in model 2 (Table 4).

In Blacks, neonatal adiposity was only predicted by the IADPSG criteria, whereas primary CS and neonatal hyperinsulinemia were only predicted by the NICE criteria, after adjustment for confounders in model 2. On the other hand, the IADPSG criteria also performed better in predicting LGA infants, based on the AIC value (Supplementary Table 5).

CONCLUSIONS

Our study compared the prevalence of GDM and performance in predicting pregnancy outcomes between the IADPSG and NICE criteria by using a subcohort of the HAPO study cohort and incorporating subgroup analysis by race/ethnicity (i.e., Whites, Hispanics, Asians [majority Chinese], and Blacks). The HAPO database provided a unique opportunity to examine adverse pregnancy outcomes independent of any intervention. Consistent with previous studies, we found that IADPSG criteria not only diagnosed more women as having GDM but also better predicted hypertensive disorders in pregnancy, primary CS, and LGA infants (10–12,18). In addition, we also found that IADPSG criteria performed better in detecting higher neonatal adiposity, hyperinsulinemia, and clinical hypoglycemia

Table 3—Associations of GDM diagnosed by the IADPSG and NICE criteria with adverse pregnancy outcomes in Hispanics, whose OGTT remained blinded (n = 1,611)

	Adjusted																				
	Unadjusted						Model 1						Model 2								
	IADPSG		NICE		OR		IADPSG		NICE		OR		IADPSG		NICE		OR				
	OR (95% CI)	AIC	OR (95% CI)	AIC	OR (95% CI)	AIC	OR (95% CI)	AIC	OR (95% CI)	AIC	OR (95% CI)	AIC	OR (95% CI)	AIC	OR (95% CI)	AIC	OR (95% CI)	AIC			
Maternal adverse outcomes																					
GH/PE	1.72 (1.32–2.24)\$	1,662.0	1.41 (0.99–2.02)	1,674.1	1.81 (1.37–2.39)\$	1,606.9	1.49 (1.03–2.16)*	1,619.9	1.43 (1.07–1.92)*	1,523.2	1.43 (1.07–2.10)	1,523.2	1.43 (1.07–2.10)	1,523.2	1.43 (0.97–2.10)	1,525.8	1.29 (0.86–1.89)	907.3	0.96 (0.55–1.59)	908.8	
Primary CS	1.74 (1.22–2.47)‡	1,011.3	1.38 (0.84–2.19)	1,018.7	1.45 (0.99–2.10)	917.7	1.07 (0.62–1.76)	921.2	1.29 (0.86–1.89)	907.3	1.29 (0.86–1.89)	907.3	1.29 (0.86–1.89)	907.3	1.29 (0.86–1.89)	908.8	1.29 (0.86–1.89)	907.3	1.29 (0.86–1.89)	907.3	1.29 (0.86–1.89)
Neonatal outcomes																					
LGA	2.33 (1.64–3.29)\$	989.2	1.78 (1.11–2.77)*	1,004.8	2.35 (1.64–3.35)\$	993.9	1.77 (1.10–2.78)*	1,009.4	2.01 (1.38–2.90)\$	971.1	1.88 (1.14–2.99)†	977.9	1.88 (1.14–2.99)†	977.9	1.88 (1.14–2.99)†	977.9	1.88 (1.14–2.99)†	977.9	1.88 (1.14–2.99)†	977.9	1.88 (1.14–2.99)†
Adiposity	1.94 (1.29–2.89)‡	780.3	1.13 (0.60–1.98)	789.8	2.04 (1.34–3.07)\$	784.2	1.18 (0.62–2.08)	794.7	1.62 (1.04–2.50)*	765.8	1.20 (0.62–2.16)	770.0	1.20 (0.62–2.16)	770.0	1.20 (0.62–2.16)	770.0	1.20 (0.62–2.16)	770.0	1.20 (0.62–2.16)	770.0	1.20 (0.62–2.16)
Clinical hypoglycemia	1.09 (0.05–8.52)	59.9	—	—	0.95 (0.05–7.98)	61.0	—	—	0.79 (0.04–7.22)	71.0	—	—	—	—	—	—	—	—	—	—	—
Hyperinsulinemia	2.34 (1.63–3.33)\$	927.8	1.81 (1.12–2.85)*	942.6	2.18 (1.51–3.14)\$	927.3	1.65 (1.01–2.62)*	939.9	1.98 (1.35–2.90)\$	922.5	1.52 (0.92–2.43)	931.7	1.52 (0.92–2.43)	931.7	1.52 (0.92–2.43)	931.7	1.52 (0.92–2.43)	931.7	1.52 (0.92–2.43)	931.7	1.52 (0.92–2.43)

Model 1 was adjusted for maternal age, parity, ethnicity, and study center. Model 2 was adjusted for covariates in model 1 plus maternal height, BMI, and gestational age at OGTT, smoking status, alcohol use, maternal urinary tract infection (for GH/PE only), hospitalization before delivery, and mean arterial pressure at OGTT (except for GH/PE), family history of diabetes, and baby's sex. OR, odds ratio; — inadequate number of patients for analysis. *P < 0.05; †P < 0.01; ‡P < 0.001; §P < 0.0001.

than the NICE criteria. The better prediction of IADPSG on primary CS, LGA, and neonatal adiposity remained significant after adjustment for all potential confounders.

In Hispanics, we found that IADPSG criteria diagnosed more women with GDM, but at the same time, performed better in detecting hypertensive disorder, primary CS, LGA, and neonatal adiposity and hyperinsulinemia than did the NICE criteria. Moreover, our results also suggested that NICE criteria could not predict the risk of hypertensive disorders, primary CS, and neonatal hyperinsulinemia and adiposity compared with the IADPSG criteria. The better performance of IADPSG persisted after adjustment for multiple potential confounders. Of note, a previous study in a Spanish population suggested that IADPSG criteria also could be more cost-effective by saving €143 per woman, when IADPSG was implemented instead of Carpenter-Coustan criteria in the diagnosis of GDM (19). However, our study has the limitation of not being able to provide a similar cost-effectiveness analysis.

In Whites, the IADPSG criteria diagnosed more women with GDM and performed better in predicting infants with LGA and neonatal adiposity than did the NICE criteria. Nonetheless, the NICE criteria performed better in predicting hypertensive disorders in pregnancy, primary CS, and clinical neonatal hypoglycemia than did the IADPSG criteria. This finding was still observed after further adjustment for all the confounders. Thus, the choice between IADPSG and NICE criteria in Whites should be further assessed, probably with the addition of a formal cost-effectiveness analysis to balance the importance of these discordant outcomes. A cost-effectiveness analysis was conducted in a cohort including four HAPO centers from U.K. and Australia. When NICE-defined risk factors for GDM were used, NICE 2015 criteria had a net monetary benefit of £239,902 at a cost-effectiveness threshold of £30,000 per QALY and a 51.5% probability of being cost-effective, whereas the IADPSG diagnostic criteria had a net monetary benefit of £186,675 and a 27.6% probability of being cost-effective (20). However, the neonatal outcomes that were predicted differentially by the two diagnostic criteria in the current study, such as infants with LGA and neonatal adiposity, were

Table 4—Associations of GDM diagnosed by the IADPSG and NICE criteria with adverse pregnancy outcomes in Asians_{||} whose OGTT remained blinded (n = 1,695)

	Adjusted												
	Unadjusted				Model 1				Model 2				
	IADPSG OR (95% CI)	AIC	NICE OR (95% CI)	AIC	IADPSG OR (95% CI)	AIC	NICE OR (95% CI)	AIC	IADPSG OR (95% CI)	AIC	NICE OR (95% CI)	AIC	
Maternal adverse outcomes													
GH/PE	1.60 (1.06–2.41)*	1,079.2	1.12 (0.73–1.71)	1,083.6	1.36 (0.88–2.10)	1,029.7	1.04 (0.66–1.61)	1,031.5	1.11 (0.70–1.76)	1,003.6	0.99 (0.63–1.60)	1,003.8	
Primary CS	1.26 (0.89–1.75)	1,621.5	1.16 (0.84–1.60)	1,622.3	1.28 (0.89–1.82)	1,537.8	1.17 (0.83–1.63)	1,538.8	1.06 (0.72–1.53)	1,510.1	0.98 (0.69–1.39)	1,510.2	
Neonatal outcomes													
LGA	1.80 (1.19–2.65)†	1,059.3	1.77 (1.19–2.57)†	1,059.2	1.70 (1.11–2.54)*	1,060.4	1.63 (1.09–2.39)*	1,060.8	1.56 (1.00–2.39)*	1,026.1	1.71 (1.12–2.56)*	1,023.9	
Adiposity	1.98 (1.29–2.96)†	954.8	2.11 (1.41–3.11)§	951.4	2.13 (1.37–3.24)§	960.2	2.21 (1.47–3.30)§	957.5	1.87 (1.18–2.89)†	952.1	2.15 (1.40–3.24)§	947.2	
Clinical hypoglycemia	1.54 (0.79–2.80)	540.5	1.06 (0.52–1.98)	542.2	1.70 (0.86–3.13)	541.3	1.04 (0.50–1.98)	543.6	1.73 (0.86–3.27)	543.3	1.01 (0.48–1.96)	545.6	
Hyperinsulinemia	2.89 (1.91–4.30)§	879.2	2.95 (1.98–4.34)§	876.1	2.78 (1.82–4.20)§	886.7	2.86 (1.90–4.26)§	883.5	2.37 (1.52–3.63)§	880.6	2.51 (1.65–3.78)§	876.9	

Model 1 was adjusted for maternal age, parity, ethnicity, and study center. Model 2 was adjusted for covariates in model 1 plus maternal height, BMI, and gestational age at OGTT, smoking status, alcohol use, maternal urinary tract infection (for GH/PE only), hospitalization before delivery, and mean arterial pressure at OGTT (except for GH/PE), family history of diabetes, and baby's sex. * $P < 0.05$; † $P < 0.01$; ‡ $P < 0.005$; § $P < 0.001$. ||More than 92% of the Asians in the cohort were Chinese ethnicity.

not included in the aforementioned cost-effectiveness analysis. Furthermore, the population was confined mainly to Whites in the analysis as well as in the NICE review.

Lastly, IADPSG criteria diagnosed fewer women as having GDM among Asians, but better predicted women with hypertensive disorders in pregnancy, limited to the unadjusted model. Therefore, our results support the use of IADPSG criteria in an Asian population. However, most of the Asians in our cohort were Chinese; whether this result is generalizable to other Asian populations, such as South Asians, is uncertain.

Our results also illustrate the underlying reason for the difference in the prevalence of GDM between IADPSG and NICE criteria in Asians (majority Chinese), Hispanics, and Blacks. The difference lies in the proportion of women whose FPG levels fell between 5.1 and 5.6 mmol/L, 1-h PG levels >10.0 mmol/L, and the 2-h PG levels between 7.8 and 8.5 mmol/L. The former two contribute to a higher GDM frequency using IADPSG criteria, while the third contributes to the extra GDM by the NICE criteria. Since fewer Asians in the cohort fell into the range of FPG level between 5.1 and 5.6 mmol/L, but more fell into a 2-h PG level between 7.8 and 8.5 mmol/L, IADPSG diagnosed 2% less GDM than NICE criteria in this population. In contrast, Hispanic women more often fell into the range of FPG levels between 5.1 and 5.6 mmol/L and 1-h PG level ≥10.0 mmol/L, but less often into the range of 2-h PG levels between 7.8 and 8.5 mmol/L, resulting in a 12.1% higher rate of GDM with IADPSG than with NICE criteria in Hispanics. Similarly, a 10.5% higher rate of GDM with IADPSG than NICE criteria in Blacks was attributed to more women falling into the FPG level range between 5.1 and 5.6 mmol/L, but fewer falling into the 2-h PG level range between 7.8 and 8.5 mmol/L.

The current study has several strengths over previous studies comparing the IADPSG and NICE diagnostic criteria. Firstly, none of our participants received any treatment for hyperglycemia, and neither the women nor the managing clinicians were made aware of the OGTT results, in contrast to heterogeneous treatment criteria among the previous studies. The performance of NICE criteria in the Cambridge cohort might have been

underestimated because more women being diagnosed with GDM by the NICE criteria were treated accordingly instead of those diagnosed with the IADPSG criteria (10).

Secondly, our study constituted a multi-ethnic population that enabled subgroup comparisons by races/ethnicities. The review by the NICE 2015 committee as well as in other previous studies was predominantly in Whites.

Thirdly, the original HAPO study design enabled us to compare additional outcomes, including neonatal adiposity and hyperinsulinemia, which were found highly relevant to long-term metabolic health in recent studies (21,22).

Lastly, our data set also allowed the comparison of both sets of criteria after adjusting for multiple potential confounders associated with GDM and other adverse pregnancy outcomes, such as family history of diabetes and maternal BMI. Maternal BMI at midgestation is contributed to by prepregnant BMI and gestational weight gain; both maternal obesity and excessive gestational weight gain are independent risk factors for hypertensive disorders, primary CS, LGA, and neonatal adiposity (23–26).

Our study has also some limitations. For example, our data set did not have certain outcomes known to be associated with maternal hyperglycemia, namely, spontaneous preterm prelabor rupture of membranes, shoulder dystocia, neonatal jaundice, or respiratory distress syndrome.

Secondly, the Asian subgroup was mainly confined to the Chinese from the Hong Kong center; the application of IADPSG and NICE criteria to other Asian groups, such as South Asian, might be different.

Thirdly, our study did not have the information to compare the cost-effectiveness between the two sets of criteria. Nevertheless, IADPSG criteria apparently can identify fewer women as having GDM but can identify more adverse outcomes among those so identified in the Chinese population.

Lastly, HAPO study was an observational study. We still need a randomized study to address whether treatment of GDM based on IADPSG versus NICE criteria offers better pregnancy outcomes, as was suggested previously (27).

In this multiracial/ethnic population, the IADPSG criteria diagnosed more Hispanics

with GDM than the NICE criteria and performed better in the prediction of four individual adverse pregnancy outcomes in this subgroup. However, among Asians, the majority of whom were Chinese ethnicity, the IADPSG criteria diagnosed fewer women with GDM, while predicting hypertensive disorders better. Among Whites, the IADPSG criteria diagnose more women with GDM. Nonetheless, either IADPSG or NICE criteria performed better with relation to prediction of specific outcomes. One of the original goals of the HAPO study was to develop uniform standard criteria for the diagnosis of GDM using a 75-g OGTT. These criteria could then be applied worldwide based on the large number of study subjects from diverse racial/ethnic groups in order to compare incidence rates and outcomes based on various treatment strategies. Our results indicate that differences in rates and type of adverse outcomes vary depending on whether IADPSG or NICE criteria are used in different racial/ethnic groups. These differences are most likely the result of population differences in maternal genetics, metabolism, nutrition, and activity. Understanding these differences will help caregivers and organizations make informed choices about which criteria would be best applied to an individual population. Based on our results, we suggest that each country determine the diagnostic criteria for GDM based on its racial/ethnic composition as well health economics.

Acknowledgments. The authors thank the HAPO study steering committee for initiating and conducting the original study and for their kind help and support. The authors would especially like to thank Prof. A.R. Dyer (Emeritus Professor, Feinberg School of Medicine, Northwestern University, Chicago, IL) for reviewing the manuscript.

Funding. The HAPO study was funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (grant no. R01-HD34242) and the National Institute of Diabetes and Digestive and Kidney Diseases (grant no. R01-HD34243) from the National Institutes of Health. The current secondary analysis was funded by the Lee Hysan Foundation Research Grant and Endowment Fund Research Grant Schemes 2020–2021.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. Y.H. performed the data analysis and drafted the manuscript. R.C.W.M., H.D.M., D.A.S., J.L., P.M.C., and W.H.T. contributed to the collection of the original HAPO study data from

individual centers. R.C.W.M. and W.H.T. conceptualized and designed the study. J.L. and P.M.C. supervised the data analysis. All authors critically revised the manuscript and approved the final version. W.H.T. is the guarantor of this work, had full access to all the data in the study, and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this work were presented as an abstract and orally at the Joint Australian Diabetes in Pregnancy Society (ADIPS) and Society of Obstetric Medicine of Australia and New Zealand (SOMANZ) 2021 Scientific Meeting, virtual meeting, 23 July 2021.

References

- World Health Organization. Diagnostic Criteria and Classification of Hyperglycaemia First Detected in Pregnancy. World Health Organization; 2013. Accessed 1 November 2021. Available from <https://apps.who.int/iris/handle/10665/85975>
- Nankervis AMH, Moses R, Ross GP, et al.; Australasian Diabetes in Pregnancy Society. ADIPS Consensus Guidelines for the Testing and Diagnosis of Gestational Diabetes Mellitus in Australia and New Zealand (modified November 2014). Accessed 1 November 2021. Available from https://www.adips.org/downloads/2014ADIPSGDMGuidelinesV18.11.2014_000.pdf
- Hod M, Kapur A, Sacks DA, 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:S173–S211
- National Institute for Health and Care Excellence (NICE). Diabetes in Pregnancy: Management of Diabetes and Its Complications from Preconception to the Postnatal Period. NICE guideline [NG3]. Published 25 February 2015. Last updated 16 December 2020. London, National Institute for Health and Care Excellence: Clinical Guidelines
- Metzger BE, Gabbe SG, Persson B, et al.; International Association of Diabetes and Pregnancy Study Groups Consensus Panel. International Association of Diabetes and Pregnancy Study Groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. *Diabetes Care* 2010;33:676–682
- Sacks DA, Hadden DR, Maresh M, et al.; HAPO Study Cooperative Research Group. Frequency of gestational diabetes mellitus at collaborating centers based on IADPSG consensus panel-recommended criteria: the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study. *Diabetes Care* 2012;35:526–528
- McIntyre HD, Jensen DM, Jensen RC, et al. Gestational diabetes mellitus: does one size fit all? A challenge to uniform worldwide diagnostic thresholds. *Diabetes Care* 2018;41:1339–1342
- Cheung NW, Moses RG. Gestational diabetes mellitus: is it time to reconsider the diagnostic criteria? *Diabetes Care* 2018;41:1337–1338
- Bhatia M, Mackillop LH, Bartlett K, et al. Clinical implications of the NICE 2015 criteria for gestational diabetes mellitus. *J Clin Med* 2018;7:E376
- Meek CL, Lewis HB, Patient C, Murphy HR, Simmons D. Diagnosis of gestational diabetes

mellitus: falling through the net. *Diabetologia* 2015;58:2003–2012

11. Koivunen S, Viljakainen M, Männistö T, et al. Pregnancy outcomes according to the definition of gestational diabetes. *PLoS One* 2020;15:e0229496
12. Todi S, Sagili H, Kamalanathan SK. Comparison of criteria of International Association of Diabetes and Pregnancy Study Groups (IADPSG) with National Institute for Health and Care Excellence (NICE) for diagnosis of gestational diabetes mellitus. *Arch Gynecol Obstet* 2020;302:47–52
13. HAPO Study Cooperative Research Group; Metzger BE, Lowe LP, Dyer AR, et al. Hyperglycemia and adverse pregnancy outcomes. *N Engl J Med* 2008;358:1991–2002
14. HAPO Study Cooperative Research Group. Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study: associations with neonatal anthropometrics. *Diabetes* 2009;58:453–459
15. Metzger BE, Persson B, Lowe LP, et al.; HAPO Study Cooperative Research Group. Hyperglycemia and adverse pregnancy outcome study: neonatal glycemia. *Pediatrics* 2010;126:e1545–e1552
16. Feig DS, Corcoy R, Jensen DM, et al.; International Association of Diabetes in Pregnancy Study Group (IADPSG) Working Group on Outcome Definitions. Diabetes in pregnancy outcomes: a systematic review and proposed codification of definitions. *Diabetes Metab Res Rev* 2015;31:680–690
17. Burnham KP, Anderson DR. Multimodel inference: understanding AIC and BIC in model selection. *Soc Methods Res* 2016;33:261–304
18. Djelmis J, Pavić M, Mulliqi Kotori V, Pavić Renar I, Ivanisevic M, Oreskovic S. Prevalence of gestational diabetes mellitus according to IADPSG and NICE criteria. *Int J Gynaecol Obstet* 2016;135:250–254
19. Duran A, Sáenz S, Torrejón MJ, et al. Introduction of IADPSG criteria for the screening and diagnosis of gestational diabetes mellitus results in improved pregnancy outcomes at a lower cost in a large cohort of pregnant women: the St. Carlos Gestational Diabetes Study. *Diabetes Care* 2014;37:2442–2450
20. Jacklin PB, Maresh MJ, Patterson CC, 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
21. Moore BF, Harrall KK, Sauder KA, Glueck DH, Dabelea D. Neonatal adiposity and childhood obesity. *Pediatrics* 2020;146:e20200737
22. Bianco ME, Kuang A, Josefson JL, et al. Hyperglycemia and Adverse Pregnancy Outcome follow-up study: newborn anthropometrics and childhood glucose metabolism [published correction appears in *Diabetologia* 2021;64:1456]. *Diabetologia* 2021;64:561–570
23. Alberico S, Montico M, Barresi V, et al. The role of gestational diabetes, pre-pregnancy body mass index and gestational weight gain on the risk of newborn macrosomia: results from a prospective multicentre study. *BMC Pregnancy Childbirth* 2014;14:23
24. Badon SE, Dyer AR; HAPO Study Cooperative Research Group. Gestational weight gain and neonatal adiposity in the Hyperglycemia and Adverse Pregnancy Outcome study-North American region. *Obesity (Silver Spring)* 2014;22:1731–1738
25. Catalano PM, McIntyre HD, Cruickshank JK, et al.; HAPO Study Cooperative Research Group. The hyperglycemia and adverse pregnancy outcome study: associations of GDM and obesity with pregnancy outcomes. *Diabetes Care* 2012;35:780–786
26. Stuebe AM, Landon MB, Lai Y, et al.; Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network, Bethesda, MD. Maternal BMI, glucose tolerance, and adverse pregnancy outcomes. *Am J Obstet Gynecol* 2012;207:62.e1–62.e7
27. Bilous RW, Jacklin PB, Maresh MJ, Sacks DA. Resolving the gestational diabetes diagnosis conundrum: the need for a randomized controlled trial of treatment. *Diabetes Care* 2021;44:858–864