Odds, risks and appropriate diagnosis of gestational diabetes

No set of criteria for diagnosing GDM could ever fully separate women at risk and not at risk of pregnancy complications

The International Association of the Diabetes and Pregnancy Study Groups (IADPSG) diagnostic process and criteria for gestational diabetes mellitus (GDM) are designed to identify women at increased risk of a range of adverse pregnancy outcomes related to maternal hyperglycaemia; in particular, excessive fetal growth and fetal hyperinsulinaemia.[1] The relationship between maternal hyperglycaemia and adverse outcomes is continuous, and more than one elevated glucose level from oral glucose tolerance testing equates to higher glucose exposure; however, one elevated glucose value is sufficient to impart a higher risk of pregnancy complications.[2]

We note with concern the recent article in the Journal by d'Emden, which proposes "a more statistically valid basis for diagnosing GDM".[3] It suggests that the criteria for GDM diagnosis proposed in 2010 by the IADPSG[1] would result in up to 50% of women with a single elevated oral glucose tolerance test result being "inappropriately diagnosed with GDM as they do not meet the agreed risk threshold". We consider these statements to be incorrect and offer the following arguments in rebuttal.

Interpreting odds ratios and confidence intervals

No distinction appears to have been made between "odds" and "risks" in the statistical analysis. This is incorrect[4,5] and leads to a variety of erroneous conclusions.

Underlying the arguments in the article is an incorrect interpretation of odds ratios and their 95% confidence intervals (CIs). Odds ratios vary on a non-linear scale from zero to infinity, with 1.0 being the central value (point of no difference). If the 95% CI of an odds ratio does not cross unity, the association described is considered significant at the 5% level.[4,5]

It is wrong to suggest that 50% of subjects actually carry a risk lower than the reported point estimate. Rather, the confidence limits mean that, were a study to be repeated 1000 times, the actual odds ratio obtained would be expected to be less than the reported lower 95% confidence limit on 25 occasions.

D'Emden proposed that the lower 95% confidence limit should be used in place of the point estimate of the odds ratio in the determination of diagnostic thresholds for GDM. We strongly contend that the lower 95% CI is intrinsically less likely to be a true reflection of the prevailing odds. No statistical support was quoted by d'Emden, and such a method does not appear to be accepted in the literature. In practical terms, had the IADPSG taken this approach, it would have given a similar result to selecting a higher odds ratio threshold,

Summary

- The International Association of the Diabetes and Pregnancy Study Groups (IADPSG) diagnostic process and criteria for gestational diabetes mellitus (GDM) have been recommended by the World Health Organization for adoption and were widely introduced into clinical practice in Australia from January 2015 — in Queensland, the Australian Capital Territory and variably across other states.

- The IADPSG criteria identify women at increased risk of a range of adverse pregnancy outcomes related to maternal hyperglycaemia. The relationship between maternal hyperglycaemia and adverse outcomes is continuous; however, one elevated glucose value is sufficient to impart a higher risk of pregnancy complications.

- We outline the background and statistical foundations of the IADPSG approach and refute the inference that invalid statistical reasoning underlies the IADPSG approach.

- The prevalence of GDM diagnosed by IADPSG criteria may be higher or lower than with other criteria, depending on the underlying population prevalence of fasting and post-glucose load hyperglycaemia, which in turn vary with ethnicity. Studies comparing previous Australian criteria to the IADPSG criteria suggest GDM prevalence may decrease or may increase by up to 35% in specific populations with the planned change in criteria.

- Pregnancy complications have multiple potential underlying causes. No set of glucose criteria will ever be able to fully separate women and babies at risk of pregnancy complications from those who are not.

based on the conventional point estimate. The IADPSG specifically chose a point estimate of 1.75 for the determination of thresholds. The consensus process specifically considered and eventually rejected alternative prespecified odds ratio thresholds of 1.5 and 2.0.

Understanding the comparisons made using data

We believe that d'Emden misinterpreted the data that he presented in Box 2.[3] The odds ratios shown there, extracted from an article by two of us (A R D and B E M),[2] relate to a comparison of women diagnosed post hoc as having GDM by IADPSG criteria according to various combinations of oral glucose tolerance test results ≥ threshold (1, 2 or 3 abnormal values) versus women classified post hoc as not having GDM. They bear no relation to the continuous logistic regression model that was used to determine the IADPSG-recommended diagnostic thresholds. The comparison made is entirely different.

H David McIntyre
MD, FRACP[1]

Alan R Dyer
PhD[2]

Boyd E Metzger
MD[2]

1 Mater Medical Research Institute, University of Queensland, Brisbane, QLD.

2 Northwestern University, Chicago, Ill, USA.
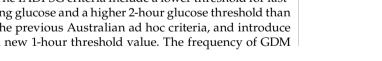
david.mcintyre@mater.org.au

For the reasons stated above, it is incorrect to say, for example, that for women with only one elevated glucose value, "50% ... did not reach the agreed risk threshold".

As previously outlined,[1] the diagnostic thresholds were defined as the average glucose values at which the odds for the three selected outcomes reached 1.75 times the estimated odds of these outcomes at mean glucose values, based on fully adjusted logistic regression models. Further, an individual woman does not have her "own" odds ratio. She, or her baby, either experience a particular outcome or they do not. As presented previously,[2] the comparison after assignment of a diagnostic label at this point should be between groups of women with and without GDM as defined by the recommended criteria.

Changing the criteria may not increase frequency of GDM diagnosis

Outside these statistical concerns, we note that many objections to the IADPSG criteria appear to be founded on a presumption of more frequent GDM diagnoses with a change to the IADPSG approach from historical criteria (in Australia, predominantly the 1991 ad hoc criteria).[6] Although an increase in the projected frequency of GDM (from 9.6% to 13.0%) in a cohort of women from Wollongong, New South Wales, was demonstrated with this change in criteria,[7] this was predominantly due to increased GDM frequency in women treated in the private sector, suggesting possible sociodemographic influences. Another Australian study, among Indigenous women from Cape York, Queensland, reported a slightly lower GDM frequency (decreased from 14.2% to 13.4%) after a change to IADPSG criteria, but noted a three-fold overall increase in GDM due to improved testing practices.[8] A cohort study from Vietnam that compared the same criteria also showed a marginally lower GDM frequency (from 20.8% to 20.4%) with a change from the ad hoc criteria to the IADPSG approach.[9]

The IADPSG criteria include a lower threshold for fasting glucose and a higher 2-hour glucose threshold than the previous Australian ad hoc criteria, and introduce a new 1-hour threshold value. The frequency of GDM detected with IADPSG criteria will inevitably vary between population groups, likely due to different underlying prevalences of fasting hyperglycaemia as opposed to post-glucose load hyperglycaemia across varying ethnicities.

The Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study

Considering the HAPO study data in particular, post hoc risk ratios for a wide range of outcomes, associated with varying diagnostic thresholds based on prespecified values (1.5, 1.75 and 2.0) of adjusted odds ratios, have been published in detail.[10] From review of these published data (summarised in the Box) it can be seen that the use of the IADPSG-recommended *odds ratio* threshold of 1.75 to determine GDM diagnostic thresholds leads to actual *risk ratios* approximating or greater than 2.0 when comparing women classified post hoc as GDM versus non-GDM (normal).

Measures selected by the IADPSG

D'Emden stated that neonatal adiposity (defined as per cent body fat >90th centile) was not a predefined outcome of the HAPO study. Although not mentioned in the primary HAPO results article,[11] it was clearly identified as a primary outcome from the earliest descriptions of HAPO study methods.[12] Full data regarding per cent body fat and cord C-peptide were available for 73% of the total HAPO cohort of 23316 neonates, representing the largest dataset of this type ever collected.

The inclusion of cord C-peptide >90th centile as one criterion used for the determination of diagnostic thresholds was criticised on the basis that "it is not a routine test performed in clinical practice". We would strongly dispute that this should be a primary reason for inclusion or exclusion of a particular measure in the definition of GDM. In fact, all the selected measures (birthweight

Risk ratios for pregnancy outcomes in women with GDM versus women without GDM*

Outcome	Non-GDM (%)	GDM (%)	Risk ratio
Birthweight >90th percentile	8.3	16.2	1.95[†]
Cord C-peptide >90th percentile	6.7	17.5	2.62[†]
Neonatal per cent body fat >90th percentile	8.5	16.6	1.96[†]
Pre-eclampsia	4.5	9.1	2.02[†]
Preterm delivery (<37 weeks' gestation)	6.4	9.4	1.47[†]
Primary caesarean section	16.8	24.4	1.45[†]
Shoulder dystocia and/or birth injury	1.3	1.8	1.44[‡]

GDM = gestational diabetes mellitus. * When using the International Association of the Diabetes and Pregnancy Study Groups recommended threshold values derived from a 75g oral glucose tolerance test for GDM diagnosis (fasting, ≥5.1mmol/L; 1-hour, ≥10.0mmol/L; 2-hour, ≥8.5mmol/L) applied to the Hyperglycemia and Adverse Pregnancy Outcome study cohort. † $P < 0.001$. ‡ $P < 0.01$. ◆

>90th centile, per cent body fat >90th centile and cord C-peptide >90th centile) are based on understanding of the pathophysiology connecting maternal hyperglycaemia to pregnancy complications. They are measures of key features of diabetic fetopathy (excess growth, excess adiposity and hyperinsulinism, respectively). Their inclusion relates directly to this nexus and we strongly contend that this is well grounded, irrespective of current clinical practice, which arguably may fall short of qualifying as a gold standard.

D'Emden also suggested that these measures were selected by the IADPSG "because they had had the greatest difference between the at-risk group and the normal group". As participants in the IADPSG consensus process, we categorically deny this assertion. By definition, any consensus process is post hoc, as it must consider information that has already been collected. Although HAPO study investigators were involved in the consensus process, they comprised a minority of the IADPSG consensus panel, which was clearly distinct from the HAPO group. Therefore, they were not constrained to the use of primary HAPO outcome variables in selection of diagnostic thresholds. We can affirm that the measures selected and the nature of further statistical calculations were decided a priori. Further, until the thresholds were selected, there were no at-risk and normal groups to compare, so this assertion is incorrect.

Conclusion

The thresholds recommended by the IADPSG, although greatly influenced by the HAPO study, also took into consideration other available epidemiological data[13-16] as well as the inclusion criteria and results of the two major randomised clinical trials in this area.[17,18] In common with any dichotomous classification of essentially continuous variables, the thresholds for GDM diagnosis are, by nature, arbitrary. Further, pregnancy complications have multiple potential underlying causes. No set of criteria for diagnosing GDM could ever fully separate women at risk and not at risk of pregnancy complications. However, the IADPSG criteria do represent a considered consensus view from an expert group and are underpinned by high-level statistical expertise.

The World Health Organization agreed to adopt the IADPSG criteria,[19] which have been accepted for local implementation in Australia by a consensus group including the Royal Australian and New Zealand College of Obstetricians and Gynaecologists, the Australian Diabetes Society, Royal College of Pathologists of Australia and the Australasian Diabetes in Pregnancy Society. From January 2015, the criteria were introduced into clinical practice in Queensland, the Australian Capital Territory and variably across other states. Other local groups such as the Royal Australian College of General Practitioners have dissented, making complete local agreement unlikely. However, we continue to support the underlying methods used to develop the IADPSG consensus criteria and contend that they represent a well reasoned approach to the diagnosis of GDM.

References are available online at www.mja.com.au.

1. 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.

2. Metzger BE, Dyer AR. Comment on d'Emden. Do the new threshold levels for the diagnosis of gestational diabetes mellitus correctly identify women at risk? Diabetes Care 2014; 37: e30. *Diabetes Care* 2014; 37: e43-e44.

3. d'Emden MC. Reassessment of the new diagnostic thresholds for gestational diabetes mellitus: an opportunity for improvement. *Med J Aust* 2014; 201: 209-211.

4. Cockburn DM. How to make clinical decisions from statistics. *Clin Exp Optom* 2006; 89: 176-183.

5. Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE, editors. Regression methods in biostatistics: linear, logistic, survival, and repeated measures models. First edition. New York: Springer, 2005.

6. Martin FI. The diagnosis of gestational diabetes. Ad Hoc Working Party. *Med J Aust* 1991; 155: 112.

7. Moses RG, Morris GJ, Petocz P, et al. The impact of potential new diagnostic criteria on the prevalence of gestational diabetes mellitus in Australia. *Med J Aust* 2011; 194: 338-340.

8. Davis B, McLean A, Sinha AK, Falhammar H. A threefold increase in gestational diabetes over two years: review of screening practices and pregnancy outcomes in Indigenous women of Cape York, Australia. *Aust N Z J Obstet Gynaecol* 2013; 53: 363-368.

9. Tran TS, Hirst JE, Do MA, et al. Early prediction of gestational diabetes mellitus in Vietnam: clinical impact of currently recommended diagnostic criteria. *Diabetes Care* 2013; 36: 618-624.

10. Metzger BE, Gabbe SG, Persson B, et al; International Association of Diabetes and Pregnancy Study Groups (IADPSG) Consensus Panel Writing Group; Hyperglycemia and the Adverse Pregnancy Outcome (HAPO) Study Steering Committee. The diagnosis of gestational diabetes mellitus: new paradigms or status quo? *J Matern Fetal Neonatal Med* 2012; 25: 2654-2659.

11. 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.

12. HAPO Study Cooperative Research Group. The Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study. *Int J Gynaecol Obstet* 2002; 78: 69-77.

13. Sacks DA, Greenspoon JS, Abu-Fadil S, et al. Toward universal criteria for gestational diabetes: the 75-gram glucose tolerance test in pregnancy. *Am J Obstet Gynecol* 1995; 172: 607-614.

14. Sermer M, Naylor CD, Farine D, et al. The Toronto Tri-Hospital Gestational Diabetes Project. A preliminary review. *Diabetes Care* 1998; 21 Suppl 2: B33-B42.

15. Jensen DM, Damm P, Sørensen B, et al. Clinical impact of mild carbohydrate intolerance in pregnancy: a study of 2904 nondiabetic Danish women with risk factors for gestational diabetes mellitus. *Am J Obstet Gynecol* 2001; 185: 413-419.

16. Schmidt MI, Duncan BB, Reichelt AJ, et al. Gestational diabetes mellitus diagnosed with a 2-h 75-g oral glucose tolerance test and adverse pregnancy outcomes. *Diabetes Care* 2001; 24: 1151-1155.

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

18. Landon MB, Spong CY, Thom E, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. A multicenter, randomized trial of treatment for mild gestational diabetes. *N Engl J Med* 2009; 361: 1339-1348.

19. World Health Organization. Diagnostic criteria and classification of hyperglycaemia first detected in pregnancy. Geneva: WHO Press; 2013. ∎