Perspectives on the Proposed Gestational Diabetes Mellitus Diagnostic Criteria

Oded Langer, MD, PhD, Jason G. Umans, MD, PhD, and Menachem Miodovnik, MD

To date, The International Association of the Diabetes and Pregnancy Study Groups (IADPSG) criteria for the diagnosis of gestational diabetes mellitus (GDM) have not been analyzed systematically for medical, social, and economic ramifications if used in substitution for the current GDM diagnostic criteria. The IADPSG dependence on expert opinion and consensus rather than on rigorously obtained outcome measures is concerning given the dramatic changes in clinical intervention and medical-resource reallocation that would follow their wide adoption. This commentary attempts to highlight needed research as well as the key knowledge gaps that should prevent adoption of the revised criteria until their effect on perinatal outcomes and health care costs is determined. In light of the overall, ethnic, and regional variation in GDM prevalence and the demands of increased GDM diagnosis on clinical resources, it may not be realistic and practical to impose universal strategies and standards for diagnosis. The newly proposed criteria may affect medical care negatively, unnecessarily stigmatize patients with a “sick label,” and adversely affect health care costs without ensuring the desired improvements in maternal and neonatal outcomes. This commentary serves as a caution to not promote a new endeavor until it has been compared rigorously with current practice and its implications are understood fully. (Obstet Gynecol 2013;0:1–6)

DOI: http://10.1097/AOG.0b013e31827711e5

In much of medical practice, we are unsure as to what really works. Even though diagnostic criteria rarely cause harm directly, the interventions and treatments that result from diagnosis may lead to more harm than good. As we consider changes to criteria for the detection and diagnosis of gestational diabetes mellitus (GDM), we are presented with the rare opportunity to determine which diagnostic approaches will improve pregnancy outcomes and, at the same time, rationalize the allocation of already-overburdened health care resources. Achieving this goal will depend on assessing relevant clinical outcomes from well-designed research rather than on opinions (including this one), consensus statements, passionate reactions, or public fanfare. We then may continue to argue about who pays for what, but, maybe, we will be closer to learning what’s worth paying for.

There are major principles that may justify the revision of any diagnostic strategy or criterion. The disease should be an important health problem with a significant adverse outcome. Differing glucose thresholds for diagnosis will necessarily alter disease prevalence in a manner that dictates altered clinical care to improve outcomes. Looser diagnostic criteria, resulting in increased prevalence, will be judged, in part, by the severity of clinical outcomes. For example, thresholds resulting in a high prevalence that identify and prevent potential stillbirth will be accepted by most clinicians, whereas thresholds that increase the sensitivity for prediction of large-for-gestational-age (LGA) neonates, when approximately 70% of such cases are unrelated to maternal diabetes, may not meet with the same acceptance.1,2 Furthermore, society’s willingness to allocate the financial and manpower resources to accommodate these changes in GDM prevalence and the interventions that it would require would reasonably vary with the severity of the outcomes that might be mitigated, with appropriate regard to overall health care costs.

The U.S. Preventive Services Task Force3 concluded in 2008 that the current evidence does not
justifying recommending GDM screening to all pregnant women. The American College of Obstetricians and Gynecologists’ continues to recommend a two-step approach, with nonfasting glucose screening of all pregnant women and an oral glucose tolerance test (OGTT) in only a small subset. The simplified “one-step” 75-g OGTT advocated by The International Association of the Diabetes and Pregnancy Study Groups (IADPSG)5 and the American Diabetes Association6 for all pregnant women, with diagnosis based on fasting, 1-hour, and 2-hour glucose determinations may not be cost effective in practice.

The diagnosis and management of GDM continues to be a focus of academic deliberation and practical uncertainty. In North America, the diagnostic criteria for GDM resulted from groundbreaking research by O’Sullivan and Mahan, using a 100-g OGTT with glucose measurements at baseline, 1 hour, 2 hours, and 3 hours. Their criteria later were modified to adjust for modern laboratory testing methods by the National Diabetes Data Group8 and by Carpenter and Coustan.9 Another widely used set of criteria, based on use of a 75-g OGTT, are advocated by the World Health Organization and used most widely outside of the United States.10 Moreover, many countries, including Canada and Australia, have developed their own diagnostic criteria for GDM. Despite attempts to agree on a single method (75-g compared with 100-g load) and a single diagnostic criterion, consensus has not yet been reached.4,6 Why has it been so challenging to reach agreement? Is it possible that the criteria established in different countries and regions reflect reasonable differences in lifestyle and culture, ethnicity, prevalence of obesity, or health-economic priorities?

Consensus implies compromise and suggests consideration of apparently opposing evidence and concerns, in which positions that initially appeared “black and white” are muted together into grey. Sometimes, however, the resulting compromise, divorced from the underlying evidence, evolves into “conventional wisdom,” which may be comfortable, convenient, and yet incorrect. Rather than being hypothesis-testing, consensus can be a platform on which to base subsequent research. Considerations to frame a constructive appraisal of altering practice to conform to the IADPSG proposal would address several distinct unknowns and areas of difference. Among these are:

1. What (individual or composite) clinical outcomes do we seek to improve by changing our diagnostic approach to GDM, and how do we weigh these outcomes?

2. Should we adopt a one-step or (sequential) two-step screening strategy?

3. Should OGTT results be based on a single abnormal value or require two abnormal values for diagnosis?

4. Should we lower the OGTT glucose thresholds for diagnosis of GDM?

5. Should we be concerned if different populations of women exhibit markedly varying GDM prevalence due to adoption of new diagnostic criteria?

6. Should we require a combined diagnosis and treatment trial to evaluate outcomes of increased GDM diagnosis rather than assume that all newly diagnosed women will benefit from (perfect and risk-free) glucose control?

7. Should we consider cost-benefit and resource-allocation implications of increased GDM prevalence as consequences of broadened diagnosis?

8. Would women newly labeled as having GDM using the proposed IADPSG criteria expect a prognosis comparable with that expected for women classified as having GDM by the current criteria?

Many have grouped these considerations together; we suggest that they each demand independent appraisal. Consensus might then lead to studies that are based on prospectively identified and clinically important outcome measures that would compare screening and testing protocols, compare diagnostic thresholds, and assess the effect of resulting changes in GDM prevalence on medical systems and finances.

To eliminate the masking effect of the addition of the subset of women now newly diagnosed as having GDM, it is imperative to evaluate the possible increased prevalence, the magnitude of adverse perinatal outcomes, and the cost of care. There is a subset of patients who are diagnosed by the suggested IADPSG criteria (OGTT of 92 mg/dL, 180 mg/dL, and 153 mg/dL) but not classified as having GDM by the current criteria (Carpenter and Coustan: OGTT 95 mg/dL, 180 mg/dL, 155 mg/dL, 140 mg/dL or National Diabetes Data Group: OGTT 105 mg/dL, 190 mg/dL, 165 mg/dL, 145 mg/dL). Lindsay et al11 report that, of 600 women with abnormal screening, 139 (23%) had one abnormal value using the Carpenter and Coustan criteria. Langer et al12 using the National Diabetes Data Group criteria in 2,461 patients with GDM, found that 812 (33%) had one abnormal value on the OGTT results. The evidence to change the diagnostic criteria (IADPSG) will need to demonstrate the presence of increased short-term outcomes not influenced by physician behavior, such as cesarean delivery, neonatal intensive care unit admission, or induction of labor. Long-term outcomes, such as
metabolic syndrome, obesity, and type 2 diabetes, also will need to be considered.

Different endpoints that evaluate outcome should be weighted differently. For example, the odds ratio to prevent perinatal mortality may be considered clinically significant even at values of 1.1 or 1.2. By contrast, for a less severe clinical endpoint (eg, LGA or perinatal death and several neonatal complications. It did, however, reduce the risks of fetal overgrowth, shoulder dystocia, cesarean delivery, and hypertensive disorders.13,14 In another study, the benefits of treatment of GDM were evaluated using the World Health Organization diagnostic criteria. Fasting plasma less than 140 mg/dL was used as the inclusion criteria.15 Patients were randomized into intervention or routine care. The authors conclude that treatment of GDM will reduce serious perinatal morbidity, with the potential to enhance women’s health. However, this study cannot be used to confirm the use of the IADPSG recommendations based on the Hyperglycemia and Adverse Pregnancy Outcome criteria because the Hyperglycemia and Adverse Pregnancy Outcome study had a different inclusion criterion of less than 105 mg/dL. By contrast, a large case–control study reports significant decreases in a composite outcome measure and in LGA neonates after treatment of patients below the current diagnostic criteria (fasting less than 95 mg/dL).16

In 2008, the large, landmark Hyperglycemia and Adverse Pregnancy Outcome study15 was conducted in 15 centers across nine countries to evaluate the association between mild hyperglycemia and adverse pregnancy outcome using the 75-g OGTT. The inclusion criteria were fasting plasma level 105 mg/dL or less or 2-hour plasma glucose level 200 mg/dL or less. Participants were blinded to the OGTT results, as were their physicians. Primary outcomes were birth weight above the 90th percentile for gestational age, primary cesarean delivery, clinically diagnosed neonatal hypoglycemia, and cord-blood serum C-peptide level above the 90th percentile. Secondary outcomes were delivery before 37 weeks of gestation, shoulder dystocia, and neonatal intensive care unit admission.

Adjusted odds ratios for the association between maternal glucose as a categorical variable and the primary outcome measures demonstrated a continuum of increasing risk with fasting, 1-hour, and 2-hour plasma glucose levels without an obvious threshold. Results for the anthropometric and C-peptide measurements suggested a link between maternal glycemia and neonatal adiposity, perhaps mediated by fetal insulin. This finding is in accord with the 50-year-old Pedersen hypothesis that maternal glucose transported to the fetus across the placenta causes fetal hyperglycemia, which in turn stimulates fetal insulin release that acts as the secondary messenger leading to fetal overgrowth.16 Associations also were observed between maternal glucose measurements and the secondary outcomes, although these tended to be weaker. Importantly, the design of the Hyperglycemia and Adverse Pregnancy Outcome study deviated from standard practice in the United States and from prevailing American Diabetes Association and American College of Obstetricians and Gynecologists recommendations4,6 by failure to include a 3-hour glucose measurement in the OGTT. As a result, the Hyperglycemia and Adverse Pregnancy Outcome study may have underestimated the occurrence of some outcome measures (eg, macrosomia). In addition, the inclusion of participants with fasting glucose levels of 105 mg/dL or less and postprandial levels of 200 mg/dL or less may have led to the inclusion of some women who would have been labeled as having GDM by the Carpenter and Coustan criteria, denying them standard treatment and yet including their morbid outcomes.
Data from the Hyperglycemia and Adverse Pregnancy Outcome study have enhanced our understanding of the association between even mild hyperglycemia and adverse perinatal outcomes. However, these data were never intended to serve as a foundation for GDM diagnostic criteria, as suggested by the IADPSG. The intended goal of the Hyperglycemia and Adverse Pregnancy Outcome Study Cooperative Research Group was to “…clarify the risks of adverse outcomes associated with various degrees of maternal glucose intolerance less severe than that in overt diabetes mellitus.” The debate is not for or against the Hyperglycemia and Adverse Pregnancy Outcome study results and their importance, but rather whether the odds ratio of 1.75, chosen arbitrarily, should be the basis for new diagnostic criteria. Ryan suggests that the number of adverse perinatal events that potentially could be prevented using the new criteria would be small, suggesting use of an odds ratio of 2.00 that would result in GDM prevalence comparable with that observed currently, with a small increase in potential complications. Long and others have argued strongly against the use of consensus standards for new criteria and emphasized the need for properly designed prospective studies that show clear evidence of benefit after the treatment interventions that would be triggered in response to increased diagnosis. In addition, it should be noted that senior Hyperglycemia and Adverse Pregnancy Outcome investigators are also the authors of the IADPSG consensus document.

In nonpregnant women, the rate of type 2 diabetes has increased significantly, paralleling the worldwide epidemic of obesity. Although there are limited data suggesting that obesity may contribute to the increasing incidence of GDM, this remains an area of active investigation. Furthermore, information derived from nonpregnant patients cannot be used as an explanation for the prevalence of GDM. The prevalence of GDM, using current diagnostic criteria, ranges from 4% to 6% in the United States. Factors contributing to variation in GDM prevalence may include obesity, racial and ethnic disparities, and the diagnostic criteria in use. In fact, studies in the past decade have demonstrated an increase in GDM prevalence related directly to the use of different diagnostic criteria. Comparing National Diabetes Data Group criteria with the Carpenter and Coustan criteria, GDM prevalence was reported by Berggren et al as 3.3% compared with 4.6%, by Gokcel et al as 4.1% compared with 6.5%, by Chou et al as 3.5% compared with 7.8%, and by Köşüş et al as 5.6% compared with 8.1%, respectively. The authors conclude that, in their target population, there was no incremental clinical benefit to the use of the Carpenter and Coustan criteria. Some authors even advocate ethnic-specific guidelines for GDM diagnosis.

The recommended IADPSG criteria would result in even more dramatic changes in apparent GDM prevalence, categorizing approximately one in five pregnant women (17.8% of the overall Hyperglycemia and Adverse Pregnancy Outcome sample) as having GDM, whereas application of these criteria to each Hyperglycemia and Adverse Pregnancy Outcome study site would have revealed wide variation (9.3–25.5%) in GDM prevalence. Recent reports have echoed this variation in prevalence, suggesting that GDM would be diagnosed in 12.4–37.3% of women in different regions and populations. The new criteria have not extended the options available for GDM diagnosis; they have, however, raised the question of whether, given regional differences, a single, uniform criterion can adequately identify GDM and at what cost.

Often ignored, the IADPSG strategy relies on reproducibility of the OGTT, which has been evaluated in multiple studies in both pregnant and nonpregnant states. Variability of OGTT results is greater in pregnant than in nonpregnant women. When values from two OGTTs were compared, fasting (differences of 7–19 mg/dL) and 1-hour plasma glucose levels (differences of 30–45 mg/dL) varied least. It follows that lower diagnostic thresholds or the use of a single abnormal value criterion each will contribute to greater variability in OGTT results. Even when National Diabetes Data Group criteria are used, OGTT results are not reproducible in 24% of cases. We do not as yet have evidence of the level of reproducibility of the IADPSG criteria.

Any decrease in the diagnostic thresholds that results in increased prevalence of GDM necessarily will lead to increased costs in both diagnosis and treatment. There have been no adequately designed and powered randomized studies to answer the question of whether new diagnostic criteria as recommended by the IADPSG consensus will be clinically, socially, and economically efficient and effective. The medical model does not exist in a cultural vacuum. The social nature of illness is particularly evident with the application of a medical label (ie, GDM). When a woman is identified as having an expensive or feared medical condition, it significantly disrupts her life and often increases her psychological stress. The experience of being labeled “sick” has both short-term and long-term social and physical consequences.

If the new IADPSG criteria (one-step diagnostic test, using one abnormal value on the OGTT) are implemented, the result will be a significant increase
in patient volume. Hospitals, clinics, and private physician offices will need to determine how to deploy and manage personnel and space. Access to enhanced testing may lead to greater demands on staff as productivity rises and more is expected of them. How many more doctors, nurses, diabetic nurse educators, and clerks will we need to accommodate the increased number of patients who will be screened and diagnosed? How will current physical facilities be expanded to perform more tests and deliver dramatically more care and counseling? Will current laboratories be able to handle the increased volume accurately and efficiently?

New diagnostic criteria should be selected with the patient in mind, minimizing the burden of testing so that it will be proportionate to expected risk. Universal OGTT screening will entail costs and burdens of missed work and complicated child care. The number of work days lost to testing multiplied by the millions of pregnant women in the United States surely will impose a large and predictable economic burden.43

Recent studies evaluating the cost-effectiveness of the one-step method recommended by the IADPSG and the two-step method currently used in many institutions worldwide suggest that the two-step approach is more cost-efficient.44–46 A single study found that the one-step approach would be more beneficial, but only when postpartum counseling and care are included. When the postpartum component was omitted, there was no outcome benefit.47

We need to press the pause button. Rather than moving forward to universal adoption of diagnostic criteria that have been tested only by an elaborate retrospective “thought experiment” using data from the Hyperglycemia and Adverse Pregnancy Outcome study, we have the moral responsibility to encourage empirical prospective research that will maximize certainty, consistency, and predictability of the clinical and social effects of this change in GDM diagnosis.

REFERENCES
Reevaluating New GDM Diagnostic Criteria


