1- or 2-step: Outcomes studied in GDM screening

Anne Paxton

June 2021—If screening for gestational diabetes mellitus were a dance competition, it might have a contest between quickstep and paso doble as its signature event. That tournament could pit the one-step testing protocol (twice as likely to diagnose GDM) against the two-step testing protocol (significantly easier for pregnant women to adhere to).

Which would prevail? In the real-life dance-off between one-step and two-step, despite helpful new input from a large-scale clinical trial, it may still be quite some time before the judges are able to declare a winner.

In Europe, consensus is strong that the one-step screening of pregnant women for gestational diabetes mellitus, which requires the patient to arrive in a fasting state, is preferable to two-step screening. But without conclusive evidence, medical opinion in the United States on this issue has long been divided. In a notable example of the disagreement, the American Diabetes Association (ADA) sides with Europe in favoring one-step while the American College of Obstetricians and Gynecologists favors the two-step approach.

The dissension about GDM screening goes beyond that, says David B. Sacks, MB, ChB, senior investigator and chief of the clinical chemistry service for the National Institutes of Health Clinical Center. From organization to organization, “People can’t agree on whether you should screen for GDM, when you should screen, how you should screen, how much glucose you should give during screening, whether the protocol should last two hours or three hours, or what your cutoff should be.”

A new study published March 11 in the *New England Journal of Medicine* doesn’t try to settle the dispute over which protocol is superior (Hillier TA, et al. *N Engl J Med.* 2021;384[10]:895–904). But the large pragmatic, randomized clinical trial of gestational diabetes screening offers findings on a related matter: Are there differences in maternal and perinatal primary outcomes depending on which screening test is used?

GDM experts agree that the findings of the study, despite its limitations, shed light on the best way to screen pregnant women for a condition that affects six percent to 25 percent of them (depending on diagnostic criteria) by increasing the risk of stillbirth, neonatal death, and multiple serious conditions in mothers and their babies.

The Kaiser study consisted of 23,792 pregnant women who received care at Kaiser Permanente Northwest or Kaiser Permanente Hawaii and were randomly assigned to one-step screening or two-step screening. (The one-step protocol entails a glucose tolerance test in which the blood glucose levels are obtained fasting and for two hours after oral administration of a 75-g glucose load. Two-step screening includes a glucose challenge test in which the blood glucose level is obtained one hour after oral administration of a 50-g glucose load in the nonfasting state, and, if the first test is positive, a three-hour oral glucose tolerance test with a 100-g glucose load in the fasting state as the second diagnostic step.)

The trial found that the single-step approach resulted in detection of GDM in twice as many women as the two-step screening, but there were no significant between-group differences in the risks of the primary outcomes relating to perinatal and maternal complications. Primary outcomes were a GDM diagnosis, hypertensive disorders of pregnancy, primary cesarean delivery, large-for-gestational-age infants, and the perinatal composite outcome.

In the same issue, Brian Casey, MD, of the University of Alabama Department of Obstetrics and Gynecology, in his editorial, writes that the perinatal benefits of the GDM diagnosis with the use of the single-step approach “appear to be insufficient to justify the associated patient and health care costs of broadening the diagnosis.” He adds: “Refocusing attention on interventions in women who are at risk for the development of diabetes is more likely to yield substantive benefits.”

There’s good evidence from earlier trials that screening for gestational diabetes in pregnancy improves perinatal

outcomes, says the lead author of the study, endocrinologist Teresa A. Hillier, MD, MS, distinguished investigator at the Kaiser Permanente Center for Health Research in Portland, Ore., and Honolulu, Hawaii. “And those trials’ data related to the two-step screening. But another study called the Hyperglycemia and Adverse Pregnancy Outcome [HAPO] trial looked at one-step testing. It found in 2008 that there was a linear relationship: The higher the glucose, the more problems with large-for-gestational-age infants and other outcomes. The study told us that C-peptide levels increased with hyperglycemia too.”

Dr. Hillier

Because of this linear relationship, influential standard-setting groups including the ADA and the International Association of the Diabetes and Pregnancy Study Groups (IADPSG) recommended a single-step approach to diagnose GDM using a fasting sample. But since outcome research had been restricted to two-step–tested patients, “there wasn’t any evidence comparing outcomes head to head with both approaches,” Dr. Hillier explains.

That led to a much more robust debate among experts at the ADA and ACOG and no resolution. “At a consensus conference in 2013, the National Institutes of Health said, ‘We can’t make a recommendation yet based on current evidence.’ Everybody was hoping that identifying milder cases would improve outcomes, but we didn’t know.”

“So what we set out to do in our study that was unique is, because the one-step and two-step were both clinically recommended test protocols, we could randomize the study as part of clinical care and waive the need for individual consent, as long as participants have the right to opt out.” That was the right design for answering the question, Dr. Hillier says, because it allowed the researchers to do the trial on the scale of the whole population of pregnant women, making it possible to scrutinize different associated outcomes in mother and baby. “So it included vulnerable groups that typically wouldn’t agree to be involved in a clinical trial, including our Medicaid population and other demographics that don’t tend to be as involved in research.”

The “pragmatic” element of the study was that although patients were assigned randomly initially to either one-step or two-step, they and their clinicians were free to change them to the other testing protocol based on clinical judgment. The institutional review boards waived the need for individual consent because both tests were clinically recommended, as long as there was an option to opt out of randomized assignment based on clinical judgment. Over the course of the trial, there was more opting-out for the two-step approach, and the research team needed to extend the duration of the trial to obtain sufficient numbers. “But that was the pragmatic part of the trial; the participants had to have the option to opt out. And we did have that adherence issue,” Dr. Hillier says.

“We designed the study to ask, ‘Are there any differences between the two groups?’ Not ‘Is one better than the other?’ because that would require a different statistical model and setup. And we found there were differences in GDM incidence, with it being twice as common with the one-step versus the two-step. But there were no differences in any of our primary outcomes: large-for-gestational age, perinatal composite outcome, hypertension, and C-section rates, or in the secondary or safety outcomes.”

“There has been hope that diagnosing twice as many women by using the one-step protocol would improve outcomes,” Dr. Hillier says. “But there had never been a head-to-head trial comparing the treatment outcomes of the two tested cohorts. It was another advantage of the trial having been done at an integrated health care system where everybody received the same treatment after the diagnosis.”

The pragmatic advantage of the two-step is that everybody first does a nonfasting step. “Only about 20 to 25 percent of women fail that and have to go on to the longer three-hour glucose tolerance test. But if they don’t

fail"—and that's the majority—"they're done." To minimize the different levels of adherence, "we went out to providers and talked to them and we started getting pretty consistent feedback that they needed to switch to the two-step protocol with some patients just to ensure screening happened. And we couldn't argue with that."

In the end, 27 percent of the women randomized to the one-step ended up having the two-step. So the level of adherence was one of the study's limitations. An additional limitation was the level of participation of racial minority populations. "Our two populations in Hawaii and the Pacific Northwest are representative of those areas, but not of the overall U.S.," Dr. Hillier says. Black and American Indian populations were underrepresented in the study.

Some aspects of the laboratory testing performed and reported in the study raise additional concerns, potentially undermining its value, says the NIH's Dr. Sacks. Part of his critique relates to different cutoffs used in the glucose testing performed at Kaiser's two study sites. "At one site, they called 130 a positive and at the other they picked a cutoff of 140. Obviously, if you pick 130 as the cutoff you can identify more people than if you pick 140. That said, the number of people who went on to the second test was different but they just combined everybody." (Dr. Hillier acknowledges there were two different regional standards for cutoffs. "Among the controversies around testing for GDM is what are the right thresholds for the one-step. Should it be 130, 135, or 140? And our two regions had different standards of care for the one-step," she says. Additionally, she points to the comment about this in the supplementary appendix published online: "Only 92 cases of GDM were diagnosed at KPNW due to a GCT in the 130–139 range, suggesting that the difference in thresholds is unlikely to affect overall results.")

Dr. Sacks

In addition, Dr. Sacks notes, 18 percent, or 165, of the women in the two-step protocol did not meet the criteria for GDM based on that protocol, but they were treated as having GDM because their fasting glucose was increased. In such a case, "you're not strictly following the protocol and it muddies the waters," he says.

Also unsettling to him is the absence of information about how the glucose was measured for the trial. "People think glucose is a perfect test and it's not even close to that," he says. In this study, "there is no mention of how they handled the glucose samples, and this is critical because you can get glycolysis." In the Hyperglycemia and Adverse Pregnancy Outcome study, which Dr. Sacks considers the best designed in terms of glucose handling, "they were very, very stringent. Everybody had to immediately place the samples on ice and spin down the blood within 30 minutes to get rid of the cells to minimize glycolysis. I have no idea what they did in the Kaiser study."

One other limitation he sees is that the study could address only short-term outcomes, which leaves out significant factors that should influence a choice between one-step and two-step. For example, he says, "The HAPO study shows that 10 or 11 years following a GDM pregnancy, there was a big increase—a doubling—of obesity in the offspring of the GDM women."

The overarching concern in any discussion of GDM screening is that the one-step protocol diagnoses a much larger number of women with GDM. "The main concern the ACOG gave in 2013 for not using the one-step protocol backed by IADPSG is that it increases by two- to maybe even threefold the number of women who are diagnosed with GDM," he says. That leads to having 18 to 20 percent of pregnant women labeled as having GDM. It was for this reason that ACOG opted to support the two-step approach of a nonfasting glucose tolerance test with a three-hour test in a fasting state if glucose concentrations are high.

But a screening test, Dr. Sacks says, should err on the side of false-positives rather than risk missing cases through

false-negatives. “You don’t want to miss anybody. So it’s better to have false-positives and send them off for a diagnostic test, which would be the second test.”

The prevalence of prediabetes and diabetes together in women of gestational age is about 24 percent in the U.S., he notes. “So 20 percent is not a high number for GDM diagnosis if you think about it in those terms.” However, ACOG allows many different versions of the two-step protocol, which is another problem, in his view.

It’s difficult to predict the impact of the Kaiser study, and there may not be enough attention to the study’s limitations, Dr. Sacks says. “*New England Journal* papers are always very, very influential and change practices often. Even papers that have flaws in them, once they are published they might run a correction but nobody will see it. I expect there will be opinion pieces in different journals addressing the study’s limitations, but they won’t get nearly as much publicity as a *New England Journal* paper.”

One important aspect of the Kaiser study, in Dr. Sacks’ view, is that by focusing on patient outcomes it improves on the original criteria for diagnosing GDM, which were from the 1960s and used two standard deviations above the mean glucose tolerance as the cutoff for women likely to develop diabetes post-pregnancy. The 2008 HAPO trial on which the IADPSG guidelines are based was a perinatal outcomes-based study.

Coincidentally, the David Sacks who is a coauthor of a “Perspectives in Care” piece published in April on “Resolving the Gestational Diabetes Diagnosis Conundrum” is a specialist in neonatology and not the same person as the NIH investigator (Bilous RW, et al. *Diabetes Care.* 2021;44[4]:858–864). But both physicians agree on the need, expressed in that *Diabetes Care* editorial, for an international, multicenter, randomized controlled trial of treatment to answer the GDM screening question.

Research like the Kaiser study is centrally important to perinatologist Amy Valent, DO, director of the Diabetes and Pregnancy Program and assistant professor of obstetrics and gynecology, Oregon Health and Science University. “I specialize in endocrinopathies that occur in pregnancy, and certainly diabetes is one of the largest endocrinopathies that we see in pregnancy. Knowing who we are helping and who we’re not helping is a topic near and dear to our hearts.”

In an AACC-sponsored debate in 2019, moderated by Dr. Sacks, Dr. Valent was designated to make the case for the two-step approach to screening. But she could just as easily have advocated for the opposite point of view, because one-step is what her institution uses in practice. She is able to see both sides of the argument.

Dr. Valent

“This is always an interesting topic, if only for the fact that we still haven’t agreed on which one we should do here in the U.S. The rest of the world has adopted the one-step approach. The U.S. is really the only remaining country that still has controversies over this.” The Kaiser study and all other countries in Europe, she says, have demonstrated with their data that the one-step approach results in considerably more GDM diagnoses.

The Kaiser study is unique in looking at so many people, Dr. Valent says. “I think that’s the strongest part of the study. The population was so large and you can look at the question in a more generalizable fashion.” Like Dr. Sacks, she stresses that measuring glucose is a very sensitive thing. “How you handle the specimen, how you order the test, the time in processing, and all of that—there are a lot of variables that go into the actual value you get from that sample of blood.” Moreover, “there is interpersonal variability. Studies have shown that even if the same person were to do the same test a week apart, it could look very different.”

The cutoffs chosen for the one-step and two-step approaches are also different, she points out. “The two-step approach was to identify pregnant women who are eventually going to develop type 2 diabetes because they were at highest risk of having adverse pregnancy outcomes and maybe we can intervene sooner. Whereas the one-step approach in the HAPO study looked at actual fetal outcomes and pregnancy outcomes like C-peptide, neonatal hypoglycemia, and what are going to be the cutoffs associated with those outcomes.”

“If you’re going to start looking at outcomes, then you have to ask the question: What’s going to influence an outcome? Well, that’s glycemic control. So with insulin versus oral agents, how well the patients’ blood sugar is controlled. And then we can get to a whole rabbit hole of where our pregnancy targets should be.” The Kaiser study, in her view, wouldn’t be needed if it were “just to answer the question of which approach is going to diagnose GDM more.” It’s needed to answer the question: “If we identify more people using the one-step approach, then how can we improve outcomes? Then you have to start looking at what actually influences outcomes.”

Adherence problems did affect the Kaiser study, Dr. Valent says. “A third of the participants in the one-step approach converted to the two-step. So it’s difficult to make strong conclusions. At a person’s first prenatal visit, they got randomized right off the bat and providers were not blind to those randomization strategies. And there are providers who would say, ‘Well, I don’t think this patient is going to come back and fast for me, so I’m going to switch her over to the two-step approach so I can give her a sugar test right now.’ That was a huge limitation for this study because there was so much provider bias there. But that’s real life.”

Confirming the value of getting a screening test done is that the six percent of patients who avoided all testing have worse outcomes, she says. “It just goes to show that regardless of what you decide to do as a provider based on the evidence, you should do some sort of testing, whether it’s one-step or two-step, because the people who didn’t get any testing had worse outcomes.”

At the end of the day, Dr. Valent says, “I just talk to my referring providers and say, ‘However you decide to screen your patients, just at least screen them. That’s number one. And number two, you have to do what’s best for your practice. If you cannot sustain having twice the volume of GDM in your practice, then maybe the two-step approach is better for you right now.”

Although no pregnant mothers want to hear they have GDM, “if we can help them learn about healthy eating patterns and higher-quality food, that is a wonderful benefit. I oftentimes tell my patients with GDM that I think this is a blessing in disguise because you can change your life around and improve your metabolic health going into the future.”

The non-blindness of the Kaiser study, because the providers have to know which test their patient is receiving, has created controversy. A randomized controlled trial underway now, Dr. Valent says, does have the providers blinded as to which arm their patient is randomized. “So it will be interesting to see if they see similar results or not.”

There isn’t yet enough evidence to say which of the two approaches is better, she says. “It clearly doesn’t change outcomes to do one or the other. So should we be managing people differently? Should we be diagnosing earlier, since the fetal pancreas starts secreting insulin at about 11 to 12 weeks? By the time you’re diagnosed at 24 to 28 weeks, a lot of time has passed, especially if there’s been silent hyperglycemia. Is it a timing issue and not necessarily a test issue? How we screen is a little less important than how we should manage.”

“When you’re trying to compare a test to an outcome,” Dr. Valent says, “that makes it very challenging because you have a lot of variables that can’t be accounted for in this type of study,” such as quality of nutrition and medication management strategies. “You have to understand the limitations. Just because something doesn’t show a difference doesn’t mean it’s not important. You have to understand that maybe there are questions that weren’t able to be answered based on the study design.”

That said, she notes, there is legitimate concern about under-screening. Data released this year revealed that at one hospital in New York, only 12 of 97 women who met ACOG’s criteria for early GDM screening received it. “The

ADA and ACOG both recommend screening early in pregnancy for those individuals at higher risk for having type 2 diabetes, and we aren't doing that very well.

"What this study shows without controversy is that not screening is potentially harmful," Dr. Valent says. "If you don't screen, you have potentially worse outcomes. So screening is meaningful. Whether you do the one-step or the two-step, just screen. And let's move forward and address bigger questions."

For Dr. Hillier, the value of the Kaiser study is clear: It addresses a significant gap in the research and should advance screening and treatment of gestational diabetes. "This topic has been extremely controversial and very hotly debated because people care about how to best treat their patients and we've been lacking evidence. There are well-intended people on both sides trying to do the right thing for their patients," she says.

"I feel very satisfied that we did a very, very diligent job. And during the trial we had to go longer than we expected because of the non-adherence that we didn't predict. But I feel that my job as a researcher is to evaluate and study the two different approaches, implement that and analyze it, and then leave it to other expert organizations to decide what they want to do with the evidence."

She hopes that the study's findings will inform the clinical guidelines of the professional organizations as well as possibly the U.S. Preventive Services Task Force. "We'll have to see how they interpret it." But what the study shows is helpful, in her view. "I think anybody involved in diagnostic testing for women who may have gestational diabetes can be reassured that with either one-step or two-step, there are no differences in outcomes for women based on the screening approach."

As for the question of when the one-step versus two-step debate will be decided once and for all, Dr. Hillier isn't willing to guess. "I wish I had a crystal ball. But it's like trying to predict when a pandemic will be over."

Anne Paxton is a writer and attorney in Seattle.