Host: Welcome to the Anesthesiology journal podcast, an audio interview of study authors and editorialists.

Dr. James P. Rathmell: Hello. I’m Jim Rathmell, Professor of Anesthesia at Harvard Medical School and one of the Executive Editors for Anesthesiology. You’re listening to an Anesthesiology podcast that we’ve designed for physicians and scientists interested in the research that appears in the journal.

Today we’re going to talk with the lead author of an original research article and an accompanying editorial view that appear in the August 2019 issue.

With us today is Dr. Lauren Glance. Dr. Glance is Professor and Vice Chair for Research in the Department of Anesthesiology at the University of Rochester School of Medicine in Rochester, New York. Dr. Glance is the first author on an article that appears in the August 2019 issue of the journal and is titled “Measuring Childbirth Outcomes Using Administrative and Birth Certificate Data.” Dr. Glance, thank you for joining us.

Dr. Laurent Glance: Hi, Jim. It’s a pleasure to speak with you and Jill Mhyre about this article.

Dr. James P. Rathmell: Also with us today is Dr. Jill Mhyre. Dr. Mhyre is Professor and Chair of the Department of Anesthesiology at the University of Arkansas for Medical Sciences College of Medicine in Little Rock, Arkansas.

Dr. Mhyre wrote an editorial view together with Dr. Brian Bateman that accompanies Dr. Glance’s research article in the August 2019 issue of the journal and it’s titled “Measuring What Matters to Moms Most.” Dr. Mhyre, thank you for joining us.

Dr. Jill Mhyre: Well, thank you for inviting me. I’m excited to discuss this paper and hopefully start conversations among our readers and listeners about how to best measure and improve quality care for childbirth.

Dr. James P. Rathmell: Dr. Glance, we’ve known for some time that maternal complications during and after childbirth demonstrate wide variation from one hospital to the next. We also know that our national reporting systems don’t integrate maternal and newborn outcomes when defining hospital obstetric quality of care.

Can you tell us why you first set out to do this study and what your original hypothesis was?

Dr. Laurent Glance: So, that’s a great question. We know that whenever we look at outcomes, there’s a tremendous variation in outcomes across hospitals and that when you see this type of variation in hospital performance, it suggests that there is an opportunity for improvement.

So, for obstetrical care, we’ve seen a doubling of the pregnancy related deaths in the United States between 1987 and 2013. One of the keys to performance, it suggests that there is an opportunity for improvement.

Once the model is estimated, we can then calculate the hospital’s performance using the PE ratio, the ratio of the predicted outcome rate to the expected outcome rate. And that PE ratio is analogous to the hospital’s O/E ratio that many of our readers are familiar with, which is used to report hospital performance when nonhierarchical modeling is used.

So, with the PE ratio, the P—or the predicted rate of severe maternal morbidity—is calculated as follows: basically you take each patient that is in a hospital and you plug in all the risk factors in the model, as well as the hospital intercept term, to calculate that particular patient’s probability of experiencing the adverse outcome.

And then you take all those patients, all their predicted probabilities, and you average them together to calculate the overall predicted severe maternal morbidity rate for that particular hospital. So, now you’ve got the P, the numerator.

And then E—the denominator in a PE ratio—represents the expected severe maternal morbidity rate for the patients if they had delivered in the average hospital. So, you calculate that by using the same equation, except that you leave out the hospital intercept term; you essentially set it equal to zero because we’re assuming that the hospital has no impact on outcomes.

And so, then you calculate a confidence interval around that point estimate for each hospital PE ratio and if that PE ratio is significantly greater than 1, meaning that the hospital’s rate of severe maternal morbidity is greater than the rate of severe maternal morbidity for the hospital’s patients had they delivered at an average hospital, then that hospital is identified as a low-performance hospital; whereas, hospitals with a PE ratio significantly less than 1 are identified as high-performance hospitals.

We then use the same approach to calculate the hospital PE ratio of severe newborn morbidity. And then you take those two ratios, the PE ratio for severe maternal morbidity and the PE ratio for severe newborn morbidity, and you take what’s called the geometric mean of those two ratios and then you have essentially a PE ratio for the composite outcome that incorporates both the newborn and the maternal outcomes.

Dr. James P. Rathmell: So, I get it. You have this very large data set that you use to come up with an average outcome adjusted for the severity, all those predictive factors of poor outcome, and then you compare each individual hospital against the entire group of hospitals, that’s how you get this relative risk of having that poor outcome at each individual hospital. Is that an everyday way of understanding what you’re doing?

Dr. Laurent Glance: I think so.

Dr. James P. Rathmell: Your study was based on almost 900,000 obstetric deliveries in 2011 and 2012 and you found that the rate of severe maternal morbidity was 1.53% and the rate of severe newborn morbidity was 3.67%
across all of these California hospitals. What did you learn about the variation from one hospital to the next?

Dr. Laurent Glance: So, what we found is that hospitals that were identified as low-performance outliers in the data from 2011 had a severe maternal morbidity rate of about 2.8% compared to about 0.9% for the high-performance outliers. So, that’s about a threefold difference in complication rates for the moms.

And then similarly, hospitals that were identified as low performance for severe newborn morbidity had severe newborn morbidity rates of about 5.7% versus about 1.98% for high-performance outliers. So, again, about a threefold difference in rates.

So, when you see this kind of variation, it represents really an important opportunity for improvement if you can identify the best practices in high-performance hospitals and then transport them to lower-performing hospitals.

Dr. James P. Rathmell: That’s just an alarming degree of variation among hospitals. What conclusions did you reach and how would you envision this sort of analysis being used in the future? And do you think that mothers should have access to this information today?

Dr. Laurent Glance: So, I agree. This is really a lot of variation. I think what our work best establishes is that it would be feasible to use administrative and birth certificate data to create a national report card for obstetrical care.

Ideally our approach needs to be replicated by measure developers to create an updated version based on ICD-10 codes. So, we don’t have ICD-9 codes anymore and we’ve now transitioned to ICD-10 codes.

So, MQIP, the Maternal Quality Improvement Program, has now been renamed The Birth Registry and it’s aiming to have about 10 pilot sites contributing data by the end of the year and about 100 hospitals contributing to data by the end of 2020. So, it’s still very new.

So, what we are currently having discussions is about pursuing a parallel track in which hospitals would contribute both administrative and birth certificate data and then use that as a basis for national reporting while we’re waiting to get more hospitals onboard with EMR data.

Ultimately, the goal’s going to be to have The Birth Registry, the MQIP, serve as a basis for performance improvement and for public reporting. But to do this, we need get hospitals to contribute data and then we need to replicate our measures using contemporary ICD-10 data.

Dr. James P. Rathmell: So, what are the limitations of your study and what may be done in future studies to address the limitations?

Dr. Laurent Glance: So, really, the main limitation of our measures, of any measures based on administrative data, is the quality of the data itself. So, for example, the most common adverse outcome identified by this CDC algorithm on which we base severe maternal morbidity is maternal transfusion.

So, most clinicians define a transfusion of four units or more as a threshold because ICD codes do not specify the number of units of blood that are transfused.

So, this problem could be addressed or resolved in a national registry if we had access to the number of units transfused.

Dr. James P. Rathmell: Dr. Mhyre, I want to turn to your editorial view titled “Measuring What Matters to Moms Most” that appears with the article by Dr. Glance. You and Dr. Bateman do a terrific job of putting this article in perspective.

You start by telling us the safety of maternity care in the United States has lagged behind in other countries and continues to generate a dismal record of maternal and perinatal morbidity and mortality. Can you expand on that a bit?

Dr. Jill Mhyre: Sure. So, I think the most dramatic difference in the United States compared with outcomes in other parts of the world relates to the maternal mortality rate and, as Dr. Glance said, that ratio has more than doubled since the 1980s in the US; the only developed country in the world, where maternal mortality has actually increased.

The current ratio is disputed, but it appears to be between 16 and 18 maternal deaths per 100,000 live births according to the US Centers for Disease Control and Prevention. And this ratio is about threefold higher than comparable rates as seen in Western Europe and other developed countries.

Maternal morbidity is also increased in the US and as Dr. Lance implied, it’s measured using a CDC algorithm that includes a composite of serious complications. It’s driven largely by increases in blood transfusion. And so, again, we don’t really know whether that increases in one unit or increases in massive blood transfusion. But blood transfusion has increased fourfold since the 1990s.

Nevertheless, other serious complications like myocardial infarction, renal failure, shock and hysterectomy are also increased. Part of this may be the population of women giving birth who are older with higher rates of diagnoses when they enter pregnancy. But there may be other things going on like the increasing rate of cesarean delivery.

Dr. James P. Rathmell: Dr. Glance’s analysis revealed that there are hospitals where birth outcomes for neonates are excellent, but mothers experience unexpectedly high rates of complications and the converse was true in other institutions. Was that surprising to you and how much do you explain that finding?

Dr. Jill Mhyre: Yes, I expected some variation in the relative safety for mothers and infants at different institutions, but the magnitude of difference was substantial. Hospital performance measuring risk-standardized severe maternal morbidity does not correlate with a measure for neonatal morbidity.

The interclass correlation coefficient was just 0.016. So, what explains this? And I think there may be a number of different things going on and it really will be the work of institutions to sort this out going forward. There may be differences in culture in how different provider groups, different institutions support decision making and trade-off risks between mothers and infants.

The cesarean delivery rate is known to differ substantially between institutions and it would be interesting to explore the extent to which institutional cesarean delivery rates mediate hospital performance for mothers and for neonates. And the ratio of performance measured for each group.

Alternatively, blood transfusion is the most common maternal morbidity, sepsis, respiratory complications and shock are the most common neonatal complications and there may be differences in how institutions manage these complications that don’t really correlate with the other set of complications. So, differences in propensity to transfuse blood in a given hemotocrit in mothers may have very little to do with care practices that protect babies in neonate in terms of sepsis shock or respiratory complications.

So, going forward I’m excited about the potential for institutions and for researchers to view this measure to reflect on clinical practice and drive quality improvement for both mothers and infants, to think at the institutional level and how outcomes differ for mothers and for their infants and how to harmonize efforts to improve safety for the whole family.

Dr. James P. Rathmell: What do you see as the primary limitations of this study?

Dr. Jill Mhyre: Well, as Dr. Glance said, the study relies on administrative data and although coding accuracy is generally good, it’s not perfect. Some diagnoses are not well-represented in ICD-9 and he talked about blood transfusion and the fact that it’s not specific for substantial blood transfusion.

Another example of placentia accreta which is a major source of maternal morbidity, including blood transfusion, hysterectomy, coagulopathy, but is not specifically coded in ICD-9.

Dr. James P. Rathmell: You go on to tell us excellent obstetric and perinatal care depends on the active involvement of anesthesiologists to optimize not only analgesia and anesthesia, but also antepartum delivery planning, peripartum medical management and resuscitation. How might this type of outcome measure promote better teamwork?
Dr. Jill Mhyre: Not quite yet. As Dr. Glance said, it relies on ICD-9 and will need to be translated over to ICD-10. At the same time, it relies on linked data with the birth certificate and fetal death files and those are not yet widely available in most places.

But I think it points a way forward to a measure that could be implemented relatively quickly with—compared to the EMR data that hopefully will be coming in the future. And I think it will be a useful tool for institutions to start looking at the tradeoffs that we’re potentially making between maternal and perinatal safety.

With respect to something like pay-for-performance, I think there’s a substantial amount of work to do to really understand how this measures functions in different hospital settings. We know that a risk is concentrated in tertiary centers, but there are also patterns of delivery in the United States in which social determinants of health co-locate with hospitals.

And so, there may be socially vulnerable populations or institutions for whom measurement at specific cutoffs would be to institutional funding that may worsen outcomes over time. So, if you have a very high-risk population not because of their medical problems but because of their social risks and if those social risks are not adequately captured by the risk adjustment model and if we had funding mechanisms that then penalize the centers, we ultimately may be leading to greater disparities and outcomes for mothers and babies than we otherwise would.

But that being said, I think as a tool for institutions to improve over time and I think as a tool to look at where disparities exist, it is certainly a welcome measure for our centers.

Dr. James P. Rathmell: Dr. Glance, tell us a little bit more about what comes next for you and your research group.

Dr. Laurent Glance: So, I was going to comment on something that Jill said and I think it’s really interesting and that’s about, how do you go about trying to reduce disparities, whether they be in obstetrical care or cardiovascular care or any other area of medical care?

And I think one of the areas that’s been very strongly debated has been whether or not to include the social determinants in health in your risk adjustment model. So, we chose to do that. So, we actually do adjust for maternal race and ethnicity and educational level.

The idea is that, as Dr. Mhyre noted, is that when you have hospitals that take care of large populations of socially vulnerable patients, you would expect that their outcomes would be potentially worse and that if you don’t adjust for those social determinants of health and you have value-based purchasing, then those institutions may be unfairly penalized and that’s really one of the main arguments for using those risk factors that capture social risk in risk adjustment models.

But the other side of the argument is that if those disparities that you see are due to essentially some discriminatory practices in terms of how we care for those patients, in other words, if we give lower-value care to people who are socially vulnerable, then if you include those social determinants of health in your risk adjustment models, essentially you’ll be excusing hospitals at providing worse care to those vulnerable patient populations.

So, it’s very much of an area that is very much under a debate and it’s something that the national (sounds like: polyforum) has been looking at very seriously and it’s one of the areas that I’ve been involved with at the NQF. Now, in terms of what comes next, I think that the next step is going to be to work with ACOG to achieve consensus on using ICD-based measures to assess quality of care while we’re waiting for the EMR data to be more widely available. Every hospital has to collect ICD codes, every hospital has to collect birth certificate data.

So, if we have a national registry and each hospital contributes ICD codes and birth certificate data to that national registry, then it actually would be fairly straightforward to come up with measures just like the ones that we created but using ICD-10 data instead of ICD-9 data and could use that as the basis for national reporting, which could be used for public reporting and which could be used for value-based purchasing.

I think that all of us would prefer to use clinical data, but we’ve been working—trying to do that for a number of years now and it’s a slow exercise because although the penetration of EMRs is pretty high at most higher-volume hospitals, the reality is that even though we’ve succeeded in getting some of the major vendors like EPIC and Cerner to incorporate the data elements, the MQIP data elements into their EMRs, trying to translate that into getting all the hospitals in the US—or most of the hospitals in the US—to adopt these MQIP or compelling EMRs is a big task and that’s why we’re still very much working towards just the pilot phase of trying to get this done.

So, I think that at the risk of being trite, I think perfect could be the enemy of good. I think it really is time now to more rapidly expand our capacity to measure obstetrical outcomes using ICD-based measures while we’re ramping up our collection of EMR-based data.

And I think that this particular paper should help point the way in terms of how to do that.

Dr. James P. Rathmell: Terrific. And just to clarify, ACOG is the American College of Obstetricians and Gynecologists.

I hope today’s discussion will lead many of you listening to read this new article and the editorial view that appeared in the August 2019 issue of Anesthesiology where you can learn more about comparing maternal and fetal outcomes across hospitals and how this might be used to improve the quality of care of mothers and newborns.

Dr. Jon Wanderer from Vanderbilt and I also created an infographic that appears in the same issue titled “Better Care Through Better Measurement: Toward Healthy Mothers and Newborns” where we aim to clearly explain in infographic style the study.

Dr. Glance and Mhyre, thank you for joining me today and for the terrific explanation.

Dr. Laurent Glance: Thank you.

Dr. Jill Mhyre: Thank you. I enjoyed it.

Host: You’ve been listening to the Anesthesiology Journal podcast, the official peer-reviewed journal of the American Society of Anesthesiologists. Check anesthesiology.org for an archive of this podcast and other related content.