Published in final edited form as:

Semin Perinatol. 2016 August ; 40(5): 318–320. doi:10.1053/j.semperi.2016.03.009.

What we have learned about quality measures for intrapartum obstetrical care

Jennifer L. Bailit, MD, MPH^{a,*}, William A. Grobman, MD, MBA^b, and Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal–Fetal Medicine Units Network

^aDepartments of Obstetrics and Gynecology, Case Western Reserve University, MetroHealth Medical Center, Cleveland, OH

^bFeinberg School of Medicine, Northwestern University, Chicago, IL

Abstract

One of the most important purposes of measuring the quality of care is to improve the quality of care. Without measurement, it is not possible to determine whether changes in medical care delivery are helping patients. This article reviews important concepts in the measurement of quality obstetrical care, and demonstrates how these concepts have been informed by the Assessment of Perinatal Excellence (APEX) study performed by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) Maternal–Fetal Medicine Units (MFMU) Network.

Keywords

Quality of care; quality improvement; measurement; obstetrics

What is “quality of care”? The Institute of Medicine defines it as “the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge”.¹ Essentially, quality of care exists along a continuum between the most basic care that is acceptable (i.e., before it becomes unsafe) and the best possible care. The extent to which an organization is along that continuum is a measure of their quality of care.²

Many stakeholders have a vested interest in measuring care. Insurers want to be sure that they are getting value for their dollar, consumers want to be sure that they are cared for by doctors and institutions that will maximally help them, and healthcare providers want to be able to assure that their care is of a high-level. However, although quality of care can be defined in a straightforward fashion, measuring it is tricky. Before measuring quality for a certain aspect of care, one must be sure of why that aspect is being measured and how the information will be used, as the types of measures and the data required to obtain the measure may vary accordingly.

*Corresponding author. jbailit@metrohealth.org (J.L. Bailit).

One of the most important purposes of measuring the quality of care is to improve the quality of care. Without measurement, it is not possible to determine whether changes in medical care delivery are helping patients. Improving quality of care can be done at different levels of the healthcare system. Within a hospital, providers might use measures that are useful to them and tailored to a specific project. The rigor of validation needed for measures used in this manner may be less than for measures used to compare institutions across the nation. Also, the data requirements for “in-house” quality improvement projects are more flexible because it is reasonable to do chart reviews or direct patient observation for short periods of time on a relatively small number of patients to assess a new process or program in an effort to perform rapid cycle quality improvement. As most hospitals have reasonably stable characteristics of their patient populations over the short term, it may be acceptable to assume that the population risks are not changing substantially in a short period of time, and that any changes in quality measures are due to underlying changes in the healthcare system.

Another important concept is that a single quality measure may sometimes “improve” at the expense of other healthcare measures. For example, if one measures only the frequency of third and fourth degree lacerations, one easy way to look “perfect” is to perform a cesarean delivery on everyone. However, this would not truly be improving quality of care. Thus, it is important to group measures so that they are balanced with another appropriate measure if one wishes to obtain a more holistic view. For example, complications of vaginal delivery may be grouped with measures of cesarean complications. Similarly, maternal complications may be grouped with neonatal complication measures to be sure that the balance between maternal and neonatal outcomes remains in equilibrium.

No one measure can capture the entire spectrum of care or outcomes. Because a hospital does well on one obstetrical measure does not mean that it will do well on other obstetrical measures.³ It cannot be presumed that a hospital that has a low postpartum hemorrhage rate, for example, also will have a low rate of postpartum infections. Essentially, assumptions cannot be made about the high or low quality of care in an obstetrical in-patient unit from one measure alone.

If measures are going to be used to compare hospitals, a uniform data source from which data elements can be cheaply abstracted is preferable. It is not practical for hospitals to do detailed chart reviews to obtain data. If outcome measures are used, they should account for differences in-patient risks between the hospitals (unless it can be shown that these differences in patients do not matter with regard to the outcome). Not doing so may potentially unfairly penalize hospitals that care for more complicated patients and disincentivize hospitals from caring for these patients. In most areas, for example, there is conscious and purposeful transport of patients from one hospital to another based on maternal and/or perinatal risks. Thus, a patient with a placenta accreta suspected prior to delivery may have her care transferred to a large regional hospital with access to a blood bank and interventional radiology. Thus, a large regional hospital may have more patients at higher risk for hemorrhage; if hemorrhage rates are compared without accounting for patient risk profiles, the results with respect to quality of care may be misleading. The need for risk adjustment means that data sources must have the clinical detail and accuracy to allow this adjustment to be done meaningfully and reliably. Sources of administrative data that meet

the collection and economic requirements noted above are often suboptimal with regard to the quality of the clinical data needed for adequate risk adjustment. Finally, measures used to compare quality of care between institutions must have validation so that these measures are known to be valid across all populations. For example, it is not sufficient for models to be developed and tested for validity in the same population. This type of validation requires an statistical rigor as well as large high-quality data sources. Groups such as National Quality Foundation and Leapfrog review measures and endorse those that meet the appropriate validation criteria.^{4,5}

Once measures are chosen, the question becomes whether and how measures can be used to improve quality of care. A major way that measurement is currently being used is for benchmarking and feedback. “In an audit and feedback process, an individual’s professional practice or performance is measured and then compared to professional standards or targets. In other words, their professional performance is “audited”. The results of this comparison are then provided back to the individual. The aim of this process is to encourage the individual to follow professional standards.”⁶ Often, an institution will have its rates of an outcome or process translated into a graph on which other institutions’ rates are also displayed, often anonymously. Thus, an institution can be compared to its peers without knowing which rate belongs to which peer. These types of reports require a neutral data center to keep institutions from seeing each other’s data. A Cochrane review of audit and feedback, which included 132 papers,⁶ concluded that audit and feedback is most effective when:

1. the health professionals are not performing well to start out with;
2. the person responsible for the audit and feedback is a supervisor or colleague;
3. it is provided more than once;
4. it is given both verbally and in writing; and
5. it includes clear targets and an action plan.

These techniques have been used successfully in some quality collaborative groups in obstetrics⁷. For instance, the Ohio Perinatal Quality Collaborative used the audit and feedback technique to decrease the number of non-medically indicated deliveries before 39 weeks by 60% in 1 year.⁷ The California Quality Collaborative has decreased maternal mortality in comparison to national trends by concentrating on obstetrical hemorrhage and preeclampsia. These efforts were especially effective in the African-American population in whom maternal mortality rates are substantially higher than other racial and ethnic groups.⁶

Additionally, quality measures can be used to monitor the effects of policy changes. Clark et al. compared three different policies to help decrease non-medically indicated deliveries less than 39 weeks. They found that a hard stop policy was the most effective at changing practice.⁸ By using the quality measures in this way, Clark et al. demonstrated that quality improvement can be a rapid change rather than a long process.

The NICHD MFMU Network undertook a study that was specifically designed to assess quality measures in obstetric care. In this study—named the Assessment of Perinatal

Excellence (APEX) study—the primary objectives were to (1) establish risk-adjusted models for obstetric outcomes; (2) assess hospital performance across these outcomes; and (3) determine how variation in the frequency of these outcomes was related to variation in the types of care provided within the hospitals.⁹

Data were collected between 2008 and 2011 at all 25 hospitals of the MFMU Network. The study population consisted of women who delivered within one of the participating institutions, were at least 23 weeks of gestation, and had a live fetus on admission. Data were collected on eligible patients if they delivered during the 24 hour period of randomly selected days during the study period. The medical records of all eligible women and their newborns were abstracted by trained and certified research personnel. Data recorded included demographic characteristics, details of the medical and obstetrical history, information about intrapartum and postpartum events. All data were subjected to ongoing data edits to ensure accuracy.

Five obstetric outcomes were chosen a priori to focus on based on the fact that they represented different domains of obstetric complications, were clinically meaningful, could be affected by differences in clinical care, were ascertainable from medical records, and potentially occurred with sufficient frequency to allow valid institutional comparisons. The outcomes were venous thromboembolism, severe postpartum hemorrhage, peripartum infection, severe perineal laceration, and a composite neonatal adverse outcome. The outcomes were analyzed among the 115,502 women who were eligible during the study period.⁹

Of note, venous thromboembolism was too uncommon (0.03%) to analyze further. The other four outcomes were found to vary significantly among the hospitals, and to be associated with multiple pre-existing patient factors, including maternal age, body mass index, insurance status, gestational age or birth weight, obstetric history, diabetes mellitus, and smoking. These factors were shown to be at least somewhat, but not fully predictive, of the four obstetric outcomes. Moreover, although there was a generally high concordance between a hospital's ranking based on outcome frequency with and without the patient factors taken into account (i.e., with and without risk adjustment), it was shown that a given hospital's ranking could change dramatically after risk adjustment was performed. Also, there was evidence that a hospital's ranking on one outcome correlated poorly with its outcome on the other outcomes. That is to say that just because a hospital had a low frequency of complications for one outcome (e.g., postpartum hemorrhage) did not make it necessary likely that it had a low frequency for another outcome (e.g., infection).

Just as with obstetric outcomes, the type of care experienced by patients at different hospitals varied widely as well. For example, there was a more than a 50-fold difference in the frequency of episiotomy, while the frequency of vaginal delivery ranged from 61% to 80%. Moreover, many of the types of care were associated with obstetric outcomes. Nevertheless, it could not be shown that the variation in care among institutions could explain a significant portion of any of the studied obstetric outcomes.⁹

Based on APEX, several lessons about quality measures in obstetric care can be drawn. First, the study provided evidence that multiple patient factors that exist prior to hospitalization are associated with obstetric outcomes and that it may be quite important to account for these factors when considering how “well” a hospital performs with regard to obstetric outcomes. Second, the results suggested that performance assessment of a hospital based on a single outcome measure cannot be generalized to characterize overall quality of obstetric care, and that instead, multiple markers of quality need to be assessed and reported in order to gain insight into the obstetrical quality profile of a hospital. Lastly, it provided support for the concept that limiting assessment to “process” measures, even if those process measures have been associated with an adverse outcome, may not provide clear insight into the actual outcomes at a hospital, as the hospitals that perform “best” on those process measures do not necessarily have the best risk-adjusted rates of obstetric morbidity.

Acknowledgments

The project described was supported by grants from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) [HD21410, HD27869, HD27915, HD27917, HD34116, HD34208, HD36801, HD40500, HD40512, HD40544, HD40545, HD40560, HD40485, HD53097, HD53118]. Comments and views of the authors do not necessarily represent views of the NIH.

REFERENCES

1. Institute of Medicine (U.S.). Division of Health Care Services., Lohr KN, Institute of Medicine (U.S.). Committee to Design a Strategy for Quality Review and Assurance in Medicare, United States Health Care Financing Administration. Medicare: A Strategy For Quality Assurance. Washington, D.C.: National Academy Press; Number of pages.
2. Grobman W, Bailit J. Patient safety and quality measurement in obstetric care In: Steven G, Gabbe M, Niebyl Jennifer R, MD, Galan Henry, MD, Jauniaux Eric, RM, MD, Landon Mark et al. eds. Obstetrics: Normal and Problem Pregnancies, 6th ed. Philadelphia, PA: Saunders; 2012.
3. Bailit JL, Grobman WA, McGee P, et al. Does the presence of a condition-specific obstetric protocol lead to detectable improvements in pregnancy outcomes? Am J Obstet Gynecol. 2015;213(86):e1–e6.
4. National Quality Forum; 2016: vol. 2016.
5. The Leapfrog Group; 2016: vol. 2016.
6. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev. 2012;6:CD000259.
7. The Ohio Perinatal Quality Collaborative Writing Committee. A statewide initiative to reduce inappropriate scheduled births at 36 (0/7) to 38 (6/7) weeks' gestation. Am J Obstet Gynecol. 2010;202(2):43.e1–43.e8. [PubMed: 19766977]
8. Clark SL, Frye DR, Meyers JA, et al. Reduction in elective delivery at o 39 weeks of gestation: comparative effectiveness of 3 approaches to change and the impact on neonatal intensive care admission and stillbirth. Am J Obstet Gynecol. 2010;203(5):449.e441–449.e446. [PubMed: 20619388]
9. Bailit JL, Grobman WA, Rice MM. Risk-adjusted models for adverse obstetric outcomes and variation in risk-adjusted outcomes across hospitals. Am J Obstet Gynecol. 2013;209(5): 446.e441–446.e430. [PubMed: 23891630]