

Performance Measures

Assessing Quality in Obstetrical Care: Development of Standardized Measures

Susan Mann, M.D.
Stephen Pratt, M.D.
Paul Gluck, M.D.
Peter Nielsen, M.D.
Daniel Risser, Ph.D.
Penny Greenberg, R.N.
Ronald Marcus, M.D.
Marlene Goldman, Sc.D.
David Shapiro, Ph.D.
Mark Pearlman, M.D.
Benjamin Sachs, M.B.B.S.

In recent years, patient safety and quality of medical care have received considerable attention from patients, providers, and the media.¹⁻⁴ Yet adverse obstetrical events that occur in hospitals are not uniformly captured or reported, which reflects in part an absence of validated obstetric-specific indicators of quality or patient safety. For example, the Commonwealth Fund's International Working Group on Quality Indicators, a collaboration involving Australia, Canada, New Zealand, the United Kingdom, and the United States, recently developed a set of 21 quality indicators to be used internationally. However, not one of these indicators was related to obstetrical care.⁵

The few indicators that are tracked in obstetrics are of questionable usefulness. For example, maternal or neonatal deaths, while clearly important to track, are not useful as quality indices because they occur so infrequently. Similarly, national organizations such as the Joint Commission on Accreditation of Healthcare Organizations,⁶ the National Perinatal Information Center (NPIC),⁷ and the Maryland Hospital Quality Indicator Project⁸ have used cesarean deliveries and vaginal birth after cesarean delivery (VBAC) rates as measures of quality. However, recent patient demand for elective primary cesarean deliveries⁹ and literature questioning the safety of both vaginal breech delivery¹⁰ and VBAC¹¹ have significantly changed practice and undermined these rates as useful quality measures.

Article-at-a-Glance

Background: No nationally accepted set of quality indicators exists in obstetrics. A set of 10 outcome measures and three quality improvement tools was developed as part of a study evaluating the effects of teamwork on obstetric care in 15 institutions and > 28,000 patients. Each outcome was assigned a severity weighting score.

Measures: Three new obstetrical quality improvement outcome tools were developed. The Adverse Outcome Index (AOI) is the percent of deliveries with one or more adverse events. The average AOI during the pre-implementation data collection period of the teamwork study was 9.2% (range, 5.9%–16.6%). The Weighted Adverse Outcome Score (WAOS) describes the adverse event score per delivery. It is the sum of the points assigned to cases with adverse outcomes divided by the number of deliveries. The average WAOS for the pre-implementation period was 3 points (range, 1.0–6.0). The Severity Index (SI) describes the severity of the outcomes. It is the sum of the adverse outcome scores divided by the number of deliveries with an identified adverse outcome. The average SI for the pre-implementation period was 31 points (range, 16–49).

Discussion: The outcome measures and the AOI, WAOS, and SI can be used to benchmark ongoing care within and among organizations. These tools may be useful nationally for determining quality obstetric care.

Table 1. National Quality Indicators, 2001

National Organization	Quality Indicators
Joint Commission on Accreditation of Healthcare Organizations	Vaginal birth after cesarean section for patients with previous cesarean section Inpatient neonatal mortality Third- or fourth-degree perineal lacerations
Maryland Hospital Quality Indicator Project	Cesarean section rates Readmission rates
National Perinatal Information Center	Cesarean section rates Vaginal birth after cesarean section rate Operative vaginal delivery rate Uterine rupture Length of stay Readmission rate Obstetrical quality measures—birth trauma
Agency for Healthcare Research and Quality	None
American College of Obstetricians and Gynecologists	None

Because of the lack of nationally recognized quality indicators, external agencies are creating their own definitions of quality. For instance, the Leapfrog Group, a consortium of private-industry health care purchasers, has made its own definitions of quality—none of which are related to obstetrics.¹² Similarly, insurance companies are beginning to create their own quality measures to be used for hospital reimbursement,¹³ but these measures are not necessarily validated and may have a financial bias.

In the absence of a nationally accepted set of quality indicators for obstetrics, we set out to develop one for use in a large multicenter trial designed to study the effect of a teamwork training intervention in labor and delivery units. The teamwork study was a cluster-randomized prospective controlled trial involving 15 hospitals, seven in the intervention arm receiving teamwork training and eight in the control arm. There was a baseline data collection period of two months, a three-month

training period, and a postimplementation data collection period of five months. The selection process for the process and outcome measures for the teamwork study was the same. The results of the teamwork trial and development of the process measures will be presented elsewhere.¹⁴ We describe here the development of the outcome measures and three quality improvement (QI) tools, preliminary benchmarks for these tools, and an example of their use in a quality improvement project.

Developing the Quality Measures

Consensus Development Conferences

To select quality measures, we held consensus development conferences in June 2001 and April 2002. In preparing for them, research personnel reviewed current obstetric measures from the Joint Commission,⁶ the American College of Obstetricians and Gynecologists (ACOG), the Agency for Healthcare Research and Quality (AHRQ),¹⁵ NPIC,⁷ and the Maryland Hospital Quality Indicator Project⁸ to identify potential outcome and process measures. Table 1 (left) lists the measures existing in 2001 from these organizations. In addition, the quality indicators used as part of the departmental QI efforts from Beth Israel Deaconess Medical Center (Boston) were incorporated into an initial set of potential measures and compiled for discussion by participants at the first consensus conferences.

Participants in the consensus conferences included nursing, obstetric, and anesthesia leaders from each of the hospitals (see Acknowledgments) involved in the research project and representatives from ACOG; the American Society of Anesthesiology; the Association of Women's Health Obstetric and Neonatal Nurses; the Society for Obstetric Anesthesia and Perinatology; the Armed Forces Institute of Pathology; the U.S. Navy Bureau of Medicine and Surgery; the Office of the Surgeon General, U.S. Army; and TRICARE Management Activity (the U.S. military health system).

First Consensus Conference. In June 2001, we reviewed the potential measures and identified additional candidate measures. All participants were encouraged to suggest potential measures based on supportive evidence of existing literature. Additional suggested measures could be sent anonymously via e-mail after conclusion of the conference.

Second Consensus Conference. In preparation for the second conference, research personnel reviewed relevant literature for each of the candidate measures. Each measure was also assigned dimensions of quality on the basis of definitions from the Institute of Medicine's *Quality Chasm* report.⁴ Table 2 (right) lists these dimensions and their definitions.

The following criteria were applied to each measure:

1. Was there literature support or consensus that the indicator was a measure of quality?
2. Could the measure be universally applied to different practice environments?
3. Could the measure be precisely defined?
4. Was the frequency or severity of enough significance to be included?
5. Was the measure obtainable with a reasonable amount of effort?
6. Would improved teamwork likely impact the measure?

We used literature review and the Beth Israel Deaconess Medical Center computerized quality assurance database to estimate the frequency of each of the measures for potential inclusion. Using this process, the proposed candidate measures were reduced to a final list of outcome and process measures. A workbook with precise definitions for each measure was then created.

We recognized that the frequency of each outcome measure was likely to be very low. Thus, the outcome measures were combined into an Adverse Outcome Index (AOI):

The AOI (a percentage) is defined as the number of deliveries complicated by one or more of the identified outcomes divided by the total number of deliveries. For example, if hypothetical Hospital X performed 1,000 deliveries in a year and had 93 deliveries in which the mother or infant experienced one or more of the identified outcomes, the AOI would be 9.3%.

Additional Measures

Although the AOI provides a measure of frequency of deliveries with adverse events, it does not measure the severity of these outcomes. Following the second consensus conference, a scoring system was created to measure severity of outcomes. The outcome measures in the AOI were submitted to the ACOG Quality

Table 2. Institute of Medicine Dimensions of Quality

Dimension	Definition
Safe	Avoiding injury to patients from the care that is intended to help them
Effective	Providing services based on scientific knowledge to all who could benefit and refraining from providing services to those not likely to benefit (avoiding under use and overuse, respectively)
Patient Centered	Providing care that is respectful of and responsive to individual patient preferences, needs, and values and ensuring that patient values guide all clinical decisions
Timely	Reducing waits and sometimes harmful delays for both those who receive care and those who give care
Efficient	Avoiding waste, including waste of equipment, supplies, ideas, and energy
Equitable	Providing care that does not vary in quality because of personal characteristics such as gender, ethnicity, geographic location, and socioeconomic status

Source: Institute of Medicine: *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, D.C.: National Academy Press, 2001.

Improvement and Patient Safety Committee (QuIPS). Using a consensus process, QuIPS assigned a weighted score to each measure that represented the severity of the outcome. It was predetermined that the sum of the scores of all other outcomes could not be greater than the score for a maternal death. The individual scores for the 10 outcomes are listed in Table 3 (page 500).

This severity scoring system was designed for use in two ways, which will be illustrated using the hypothetical Hospital X with 93 adverse outcomes in 1,000 deliveries in a year:

First, taking the sum of the adverse outcome scores of all events and dividing by the total number of deliveries determines a Weighted Adverse Outcome Score (WAOS). If Hospital X had 93 deliveries with adverse events, the sum of the scores of these adverse events would be divided by the 1,000 deliveries. This allows one to assess the overall significance of adverse events on the unit. Second, the Severity Index (SI) is calculated by

Table 3. Adverse Outcome Index, Institute of Medicine (IOM) Dimension, and Score of Individual Indicators*

Indicator	IOM Dimension	Points
Maternal death	Safe	750
Intrapartum or neonatal death > 2,500g*	Safe	400
Uterine rupture	Safe, Effective, Timely	100
Maternal admission to ICU	Safe, Effective	65
Birth trauma	Safe, Effective	60
Return to OR / labor & delivery	Safe, Effective	40
Admission to NICU >2,500 g & for > 24 hours†	Safe, Effective	35
APGAR < 7 at 5 minutes	Safe, Effective	25
Blood transfusion	Safe, Effective	20
3°- or 4°-perineal tear	Safe, Effective	5

* ICU, intensive care unit; OR, operating room; NICU, neonatal ICU.

† The use of 2,500 grams as a cutoff birth weight for the AOI was based on the following articles, which have used this value to identify neonates unlikely to need NICU admission simply based on weight or gestational age criteria: Rohininath T, et al.: Workload and short-term outcome of babies weighing 2,500 grams of more at birth admitted to the pediatric unit of the Rotunda Hospital. *J Matern Fetal Neonatal Med* 17:130-143, Feb. 2005; Gray J.E., McCormick M.C., Richardson D.K.: Normal birth weight intensive care unit survivors: Outcome assessment. *Pediatrics* 97:832-838, Jun. 1996, and Phillips J.B., et al.: Characteristics, mortality, and outcome of higher-birth weight infants who require intensive care. *Am J Obstet Gynecol* 149:875-879, Aug. 15, 1984.

taking the same sum of scores and dividing by the number of deliveries complicated by one or more adverse events. For hypothetical Hospital X, the sum of the adverse outcome scores would be divided by 93 for the number of adverse events. The SI measures the average severity of each delivery with an adverse event.

Data Collection

We employed two data collection techniques to determine the usability of the measures. First, the 15 hospitals in the teamwork study used a standardized data collection form for each delivery during a 10-month period (an 8-week pre-implementation baseline, a 3-month intervention period, and a 5-month post-implementation data collection). Labor and delivery personnel completed the data entry concurrently. In addition, each site had a data coordinator, who ensured data accuracy and completeness. AOI, WAOS, and SI data presented here are only from the 8-week pre-implementation period of the study.

Second, we asked NPIC to determine the AOI, WAOS, and SI for their participating hospitals using discharge coding information. We compared the data derived from

NPIC with our own prospectively collected data to assess the accuracy of the NPIC collection technique. We also used the composite NPIC data to define the range of the three QI tools.

In addition, the department of obstetrics and gynecology, Beth Israel Deaconess Medical Center (Boston), experienced a major adverse sentinel event in 2000, which led to several QI efforts and the decision to participate as the lead hospital in the teamwork trial. This event and the QI changes have been described elsewhere.¹⁶ We retrospectively reviewed the AOI, WAOS, and SI at Beth Israel Deaconess Medical Center (Boston) before and after these

changes to demonstrate how these tools could be used to track the results of these changes.

Results

Consensus Conferences

During the two consensus conferences, 47 potential process and outcome measures were identified. Applying the 6 criteria described above, the final set of 10 outcome and 12 process measures was determined. The 10 outcome measures, their designated IOM dimensions of quality, and the ACOG consensus weighted scores are listed in Table 3. Table 4 (page 501) summarizes how the scoring system might be used to determine the WAOS and SI for the hypothetical Hospital X with 1,000 deliveries and an AOI of 9.3%.

Data Collection

Data Elements. During the entire teamwork study, we collected data on 28,536 deliveries, of which 7,673 deliveries occurred during the pre-implementation phase. The data collection form had 59 data fields, providing more than 452,000 data points. Data completeness for the AOI was excellent; only 130 (1.7%) of deliveries during the

pre-implementation period (range, 0–15.2% across the 15 institutions) had to be excluded from calculation because of missing or discrepant data.

AOI. The average AOI was 9.2% (range, 5.9%–16.6%), the average of the WAOS was 3 (range, 1–7) for term deliveries and 6 (range 0–34) for deliveries at less than 37 weeks, and the average SI was 25 (range, 10–39) for term deliveries and 75 (range, 20–297) for deliveries at less than 37 weeks (Table 3).

The AOI from the entire study data was strongly influenced by NICU admissions and rates of third- and fourth-degree lacerations. These two outcomes had an

average prevalence of 4.3% and 4.7%, respectively. Thus, a minority of patients who contributed to the AOI had neither of these outcomes. In addition, there was inconsistency in the criteria for NICU admits. In part, this finding was due to variations in terminology. There were 175 separate indications for admission to an NICU in the study, but many of these were overlapping (for example, “RDS” (respiratory distress syndrome) and “RDS, Tachypnea” were separate indications). The study did not have predetermined indications for NICU admission.

Overall, NPIC gathered data on 224,661 deliveries from 49 hospitals between July 1, 2004 and June 30, 2005. NPIC was able to gather all 10 elements of the AOI directly from discharge coding data in 26 (53.1%) of these hospitals. Blood transfusion data were missing from 13 (26.5%) hospitals, five-minute APGAR scores from 14 (28.5%), and maternal ICU admission data from 3 (6.1%). These data can generally be obtained from billing or other data on additional request from the medical centers. The NPIC collection process was able to identify 20 hospitals for which all 10 data elements of the AOI were

Table 4. Hypothetical Calculation for a Hospital with 1,000 Deliveries and 93 Complicated by an Adverse Event*

Outcome	Number	Points/Event	Total Points for Event
Maternal death	0	750	0
Intrapartum or neonatal death >2,500g	1	400	400
Uterine rupture	2	100	200
Maternal admission to ICU	3	65	195
Birth trauma	5	60	300
Return to OR / labor and delivery	5	40	200
Admission to NICU > 2,500g and for > 24 hours	8	35	280
APGAR < 7 at 5 minutes	12	25	300
Blood transfusion	15	20	300
3°- or 4°-perineal tear	70	5	350
Total deliveries with event	93		
Total Points			2,525
WAOS (Total points/all deliveries)			2,525/1,000 = 2.53
SI (total points/deliveries with event)			2,525/93 = 27.15

* OR, operating room; NICU, neonatal intensive care unit; WAOS, Weighted Adverse Outcome Score; SI, severity index. The total number of events is more than 93 because some deliveries might be complicated by more than one event.

available from the past six years. This included discharge data on more than 80,000 deliveries per year for the years 1999–2004. The AOI ranged between 4.5% and 25.8%, with annual averages between 9% and 11.5% for the six-year period (Figure 1, page 502). The WAOS ranged from 0.95 to 13.7, with annual averages for the 20 hospitals of 3.1–3.5 for the same years. The SI ranged from 8.5 to 59.3, with annual averages of 29.3–34.5 (Table 5, page 503).

As in the study data, NICU admissions and perineal lacerations accounted for a large percentage of the outcomes. Because of concerns about variations in definitions of these two measures raised by the study data, we recalculated the three scores without NICU admissions or perineal tears. Excluding the NICU admissions from the 2002–2004 data decreased the average AOI and WAOS by 22%–44.4% and 44%–47.6%, respectively. Similarly, excluding the perineal tears decreased these indices by 31.5%–33.3% and 3.5%–5.4%, respectively. Excluding both decreased the AOI by 66.7%–71.6% and the WAOS by 50.6%–54.7%; however, it increased the SI by 69.4%–103%.

AOI Scores for National Perinatal Center Hospitals, 1999–2004

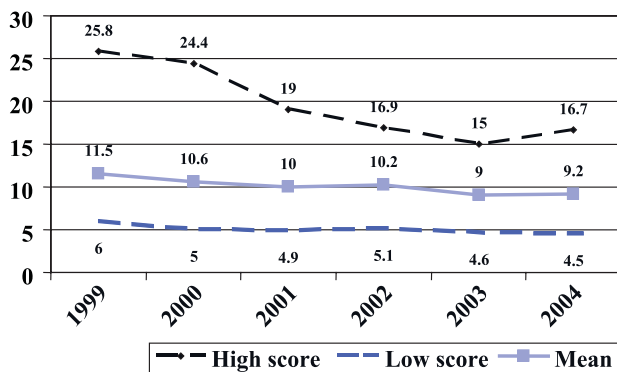


Figure 1. The AOI ranged between 4.5% and 25.8%, with annual averages between 9% and 11.5% for the six-year period. AOI, adverse outcome index.

We were able to obtain AOI data from both our own database and NPIC for the years 1999–2004. In addition, we had our internal data from 2005. We found a decrease in all three parameters after the implementation of the QI projects and the team training process. Specifically, the AOI decreased 17.2%, from a 3-year average of 5.8% in 1999–2001 to 4.8% in 2003–2005 (Figure 2, page 504). The WAOS and SI for each year demonstrated similar decreases. The AOI determined from the NPIC data was very close to our internal, prospectively gathered results, with most years being within 2% to 7% margin of error. The largest difference between the NPIC data and ours occurred in 2004 when the NPIC data was 11.5% less than ours. The accuracy of the NPIC data for the WAOS and SI demonstrated similar results. At our institution, the AOI minus NICU admits and lacerations decreased 32% from a 3-year average of 1.49% in 1999–2001 to 1.0% in 2003–2005. Notably, deliveries of less than 37 weeks gestation showed a 41% decrease in the AOI for the same periods (Figure 2).

Discussion

To improve quality of care, health care workers must first be able to clearly define and measure quality. Cesarean section and VBAC rate are currently the most universally accepted quality indicators in obstetrics.^{6–8} The use of these measures is driven by a desire to decrease cost and maternal morbidity associated with

cesarean delivery and because they are clearly defined and easily collectable measures. Recent data questioning the safety of both vaginal breech delivery and VBAC delivery and the rising demand for elective primary cesarean section have affected these measures.^{9–11} Payers are also applying their own measures of quality, which often relate more to the cost of care than to the quality of care provided.¹³ Such measures are important to payers of health care, but consumers and health care providers are more interested in and are better served by measures of care that assess the frequency of important outcomes. Such measures would be particularly important if they could discriminate among good, average, and below-average quality of care. This ability to measure the quality of care could assist consumers in making health care decisions and providers in focusing performance improvement efforts. We have developed a set of indicators, with clinical significance that can be easily and successfully measured and could be used to benchmark quality in labor and delivery units.

Others have attempted to develop quality measures in obstetrics. Waterstone and colleagues studied the predictors of severe obstetric morbidity.¹⁵ The outcomes studied included maternal hemorrhage, severe preeclampsia, sepsis, eclampsia, the syndrome of hemolysis elevated liver enzymes and low platelets (HELLP), and uterine rupture. They found that 588 (1.2%) of 48,865 patients in 19 maternity wards studied from March 1, 1997 to February 28, 1998, suffered one of these events during obstetric care. The authors did not describe their measure selection process, except to state that they chose outcomes that were clinically based and routinely measurable. Although these outcomes are certainly important, this narrow list of adverse maternal events does not include any adverse neonatal outcomes. Furthermore, the low prevalence makes studying the impact of QI efforts difficult. The AOI is a broader definition of adverse maternal events and includes significant adverse neonatal events. The average AOI rate of 9.2% is more likely to allow statistical analysis of improvement efforts.

Novicoff and colleagues found that 8,795 (80%) of 10,984 women suffered some adverse event during the peripartum period.¹⁷ These authors studied the predictors of 37 maternal and 27 neonatal adverse outcomes;

Table 5. Median and Range of AOI, WAOS, and SI for Teamwork Study and the 20 NPIC Hospitals*

	Team Study	1999 [†]	2000 [†]	2001 [†]	2002 [†]	2003 [†]	2004 [†]
Number of deliveries	7,673	81,658	84,735	85,577	85,317	87,378	64,500 [‡]
AOI (range)	9.2% (5.9%–16.6%) [§]	11.5% (6%–25.8%)	10.6% (5%–24.4%)	10.0% (4.9%–19%)	10.2% (5.1%–16.9%)	9.0% (4.6%–15.0%)	9.2% (4.5%–16.7%)
WAOS (range)	3.0 (1.0–7.0) [§]	3.5 (1.17–12.66)	3.34 (0.95–13.66)	3.25 (1.19–10.93)	3.3 (1.3–10.04)	3.11 (1.16–8.57)	3.11 (1.2–9.38)
SI (range)	25 (10–39) [§]	34.48 (8.52–56.05)	29.29 (12.33–56.04)	31.25 (14.04–57.41)	31.99 (22.32–59.25)	32.65 (19.56–57.93)	32.1 (18.35–56.08)

* AOI, adverse outcome index; WAOS, Weighted Adverse Outcome Score; SI, severity index.

[†] Data from the National Perinatal Information Center.

[‡] Data from first three quarters.

[§] Data from term deliveries only.

however, they did not describe the selection process for the measures, referring to them simply as “outcomes of interest.” Moreover, many outcomes were very similar (for example, six different outcomes related to perineal injury during delivery) or poorly defined (breast disorder or epidural problem). Thus, it may be impossible to use discharge data to study these outcomes. The real-time collection of so many outcomes would likely be very cumbersome. We had an excellent collection rate both in real time and from discharge data for the 10 elements required for the AOI, WAOS, and SI.

The use of a weighting system allows adjustment for the severity of adverse events. Pomposelli et al. used a similar weighting system in vascular surgery to assess postoperative complications.¹⁸ Novicoff created a weighting system for the 64 peripartum events they studied, weighting each outcome from 1 to 100 points on either a maternal or neonatal model.¹⁷ However, some of the outcomes were of questionable significance; for example, a woman who chose an elective primary cesarean section would receive two points for this designated adverse outcome. We chose measures that were clearly defined and could be linked to IOM dimensions of quality.⁴ Also, the fact that the ACOG Quality Improvement and Patient Safety Committee, composed of obstetricians and nurses knowledgeable about quality and safety measures, determined the weighting scores for our study, thus potentially adds national validity to our weighting system.

The WAOS and the SI are novel approaches for assessing adverse outcomes. The WAOS is a measure of the cumulative degree of adverse events in a given institution during a period of time. Simply looking at adverse event rates such as the AOI may not reflect the quality of care for a given institution. For instance, two similar institutions may have similar AOI rates; however, if one institution has a higher WAOS because of a relative preponderance of high-scoring outcomes, one might question the quality of care provided at that institution. Similarly, the SI allows for the measurement of the average severity of each adverse outcome. This may be a measure of caregivers’ ability to plan for, identify, respond to, and mitigate adverse events.

We believe that we have taken a significant step in generating an easily usable and well-defined set of obstetric outcomes and tools that can be used to measure the quality of care provided on labor and delivery suites. Additional research is warranted to better understand these outcomes and tools. Suggested research might include identification of factors that affect the rate of these outcomes. For example, the presence of 24-hour dedicated obstetric anesthesia coverage, hospitalist obstetric type coverage, the type of medical insurance, the level of neonatal intensive care, or other factors might all influence these outcomes. Additional research might also be directed at

AOI Scores at Beth Israel Deaconess Medical Center, Pre- and Post-Teamwork Intervention, 1999–2005

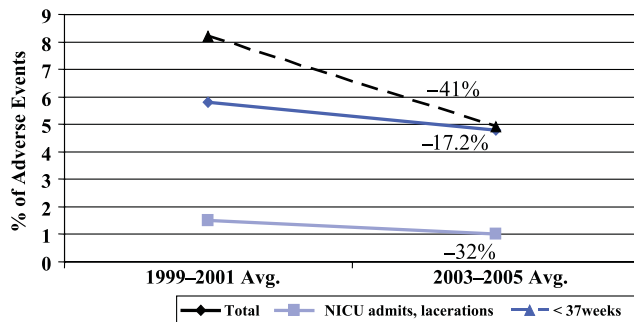


Figure 2. All three parameters decreased after the implementation of the QI projects and the team training process. AOI, adverse outcomes index.

better standardization of the indications for and descriptions of NICU admissions. This would decrease the interhospital variability in the NICU admission rate.

There are several limitations to the set of measures we described. First, we could not clearly demonstrate construct validity. The use of a consensus conference is a well recognized method for the development of quality indicators.^{19,20} The use of national experts and representatives from the governing bodies of four specialties, and strict criteria for outcome measure inclusion assured face validity of each measure and thus of the AOI. In addition, the WAOS and SI increased with high risk (< 37-week gestation) pregnancy, suggesting construct validity. Further construct validity is also suggested by our experience. We demonstrated significant improvements in all three assessment tools during a six-year period in which we were actively working to improve care through better teamwork. However, we were unable to clearly demonstrate construct validity of the AOI by comparing it with other measures of quality, in part because of the fact that there are no standard outcomes in obstetrics against which we could compare the AOI.

We were also limited in our ability to risk adjust, with the exception of preterm delivery or birthweight < 2,500 grams for NICU admissions. Currently, no case-mix adjustments are uniformly accepted in the field of

obstetrics. The ability to risk adjust can help identify which outcomes are primarily influenced by case mix and which might be preventable. Once we better understand the factors that influence the AOI, we can better risk adjust and compare hospitals. Clearly, NICU admissions and third- and fourth-degree lacerations are much more frequent than the other eight outcome measures. Expected fluctuations (due to sampling variation) in the rates of these two outcome measures could mask relatively large changes in the other outcome measures in the AOI.

In addition, the indications for NICU admissions in the teamwork study were extensive, repetitive, and poorly defined. In all, there were 175 indications for NICU admission. However, there were 39 separate categories identified for sepsis evaluation in the newborn. This suggests not only variation in the language used to identify NICU admission but also perhaps significant practice variation between the institutions regarding which newborns should go to the NICU. As NICU admission accounted for more than 33% of the total AOI score, the variation in practice may have influenced the AOI in ways unrelated to quality of care. Until national standards for NICU admission are created, it may be difficult to compare the AOI among institutions, unless the AOI is calculated without the NICU admission rate. Alternatively, a defined set of specific diagnosis-related groups for NICU term admissions could be used as a definition for this measure. Which specific DRGs to include would require additional study. In our own institution, we were able to show a downward trend in the average AOI without NICU admits and lacerations as well as in pregnancies of < 37 weeks' gestation pre- and postimplementation of the teamwork initiative. The AOI is a robust measure and is not just a reflection of lacerations and NICU admits.

Finally, the measures developed for use in this study focused on the impact of teamwork among health care providers. Thus, other measures of quality of obstetric care may exist that were not included. We specifically excluded one potential measure of quality, inadvertent injury to internal organs during cesarean delivery, because we did not feel it would be influenced by teamwork.

Conclusion

The AOI, WAOS, and SI are global measures of quality that could be used by QI specialists or obstetric chairpersons to determine the impact of QI efforts or even by health or malpractice insurers to identify best providers and best practices. Much additional work is now needed to determine what factors influence these measures, and which, if any, additional outcomes should be tracked. **1**

The American College of Obstetricians and Gynecologists does not officially endorse the Adverse Outcomes Index measure. The funding agencies assisted in the development of team work for labor and delivery. They have not been involved in the preparation of this manuscript. The views herein are those of the authors and are not to be construed as official or as reflecting the views of the Department of the Army, the Office of the Assistant Secretary of Defense, or the Department of Defense. The work described in this article was performed through a subcontract with Dynamics Research Corporation under contract to the Armed Forces Institute of Pathology and Office of the Assistant Secretary of Defense (Health Affairs)/TRICARE Management Activity Contract, Risk Management Foundation of the Harvard Medical Institutions, and Beth Israel Deaconess Medical Center Obstetrics/Gynecology Foundation.

Four of the authors are developing with Harvard's Risk Management Foundation quality assurance and team training programs that will be provided through a limited liability company, of which they are part owners.

Susan Mann, M.D., is Director of Quality Improvement, Department of Obstetrics and Gynecology, Beth Israel Deaconess Medical Center, Boston, and Stephen Pratt, M.D., is Clinical Director, Obstetric Anesthesia, Department of Anesthesia, Critical Care, and Pain Medicine. Paul Gluck, M.D., is Associate Clinical Professor, Department of Obstetrics and Gynecology, University of Miami, Miller School of Medicine, Miami. Peter Nielsen, M.D., is Chairman, Department of Obstetrics and Gynecology, Madigan Army Medical Center, Tacoma, Washington. Daniel Risser, Ph.D., is Senior Behavior Scientist, Advanced Planning and Engineering Corporation, Rockville, Maryland. Penny Greenberg, R.N., is Director of Healthcare Quality and Patient Safety, Beth Israel Deaconess Hospital, Needham, Massachusetts. Ronald Marcus, M.D., is Director, Outpatient Services Department, Department of Obstetrics and Gynecology, Beth Israel Deaconess Medical Center and Marlene Goldman, Sc.D., is Associate Professor, Department of Obstetrics, Gynecology, and Reproductive Biology. David Shapiro, Ph.D., is Senior Research Scientist, Department of Biometrics, Harvard School of Public Health, Boston. Mark Pearlman, M.D., is Associate Chief of Staff, University of Michigan Health Systems, Ann Arbor, Michigan. Benjamin Sachs, M.B.B.S., is Harold Rosenfeld Professor, Harvard Medical School, Boston. Please send requests for reprints to Susan Mann, smann@bidmc.harvard.edu.

References

1. Lockley S.W., et al. for the Harvard Work Hours, Health and Safety Group: Effect of reducing interns' weekly work hours on sleep and attention failures. *N Engl J Med* 351:1829-1837, Oct. 28, 2004.
2. Perez-Pena R.: Law to rein in hospital errors is widely abused, audit finds. *New York Times*, Sep. 29, 2004:A1.
3. Institute of Medicine: *To Err Is Human: Building a Safer Health System*. Washington, D.C.: National Academy Press, 1999.
4. Institute of Medicine: *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, D.C.: National Academy Press, 2001.
5. Hussey P.S., et al.: How does the quality of care compare in five countries? *Health Affairs (Millwood)* 23:89-99, May-Jun. 2004.
6. Joint Commission on Accreditation of Healthcare Organizations: A Comprehensive Review of Development and Testing for National Implementation of Hospital Core Measures. <http://www.jointcommission.org/NR/rdonlyres/48DFC95A-9C05-4A44-AB05-1769D5253014/0/AComprehensiveReviewofDevelopmentforCoreMeasures> (last accessed Jun. 26, 2006).
7. National Perinatal Information Center/Quality Analytic Services: *Quarterly Reports*. <http://www.npic.org/frame2.html> (last accessed Jun. 26, 2006).
8. Maryland Hospitals Quality Indicator Project: *Quality Indicator Project® Acute Care Measures*. http://www.qiproject.org/pdf/Acute_Care_Indicators.pdf (last accessed Jun. 26, 2006).
9. Minkoff H., et al.: Ethical dimensions of elective primary cesarean delivery. *Obstet Gynecol* 103:387-392, Feb. 2004.
10. Hannah M.E., et al., for the Term Breech Trial Collaborative Group: Planned caesarean section versus planned vaginal birth for breech presentation at term: A randomised multicentre trial. *Lancet* 356:1375-1383, Oct. 21, 2000.
11. Lydon-Rochelle M, et al.: Risk of uterine rupture during labor among women with a prior cesarean delivery. *N Engl J Med* 345:3-8, Jul. 5, 2001.
12. Leapfrog Group: *Leapfrog Hospital Rewards Program™. Release Notes v1.1*. 2006 Program Changes. <https://leapfrog.medstat.com/hrp/2006Changes.htm> (last accessed Jun. 26, 2006).
13. Steinbrook R. The cost of admission: Tiered co-payments for hospital use. *N Engl J Med* 350:2539-2542, Jun. 17, 2004.
14. Nielsen P., et al.: The labor and delivery teamwork intervention trial (unpublished manuscript). Madigan Army Medical Center, Tacoma, Washington, May 2006.
15. Waterstone M., Bewley S., Wolfe C.: Incidence and predictors of severe obstetric morbidity: Case-control study. *BMJ* 322:1089-1093, May 5, 2001.
16. Sachs B.P.: A 38-year-old woman with fetal loss and hysterectomy. *JAMA* 294:833-840, Aug. 17, 2005.
17. Novicoff W., Wagner D., Knauss W.: Initial development of a system-wide maternal-fetal outcomes assessment program. *Am J Obstet Gynecol* 183:291-299, Aug. 2000.
18. Pomposelli J.J., et al.: Surgical complication outcome (SCOUT) score: A new method to evaluate quality of care in vascular surgery. *J Vasc Surg* 25:1007-1014, Jun. 1997.
19. Moxey E.D., et al.: Developing a quality measurement tool and reporting format for long-term care. *Jt Comm J Qual Improv* 28:180-196, Apr. 2002.
20. McDonald K.M., et al.: *Measures of Patient Safety Based on Hospital Administrative Data—The Patient Safety Indicators*. AHRQ publication No. 02-0038, Aug. 2002. <http://www.ahrq.gov/downloads/pub/evidence/pdf/psi/psi.pdf> (last accessed Jun. 23, 2006).