Learning from Adverse Events in Obstetrics: Is a Standardized Computer Tool an Effective Strategy for Root Cause Analysis?

Beth Murray-Davis, RM, PhD,1,2 Helen McDonald, RM, MSc,2,3 Fiona Cross-Sudworth, RM,4 Rashid Ahmed, BSc,1,2 Julia Simioni, MSc,1,2 Sharon Dore, RN, PhD,1 Michael Marrin, MD,5,6 Judy DeSantis, RN,7 Nicholas Leyland, MD, MHCM,1 Jason Gardosi, MD, FRCSEd, FRCOG,4,8 Eileen Hutton, PhD,1,2,9 Sarah McDonald, MD, MSc1,9,10,11

1 Department of Obstetrics and Gynecology, McMaster University, Hamilton ON
2 Midwifery Education Program, McMaster University, Hamilton ON
3 Department of Family Medicine, McMaster University, Hamilton ON
4 West Midlands Perinatal Institute, Birmingham, United Kingdom
5 Department of Pediatrics, McMaster University, Hamilton ON
6 Division of Neonatology, McMaster Children’s Hospital, Hamilton ON
7 Division of Women’s Health, Hamilton Health Sciences, Hamilton ON
8 Department of Maternal and Perinatal Health, University of Warwick, Coventry, United Kingdom
9 Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton ON
10 Division of Maternal Fetal Medicine, McMaster University, Hamilton ON
11 Department of Radiology, McMaster University, Hamilton ON

Abstract

Objective: Adverse events occur in up to 10% of obstetric cases, and up to one half of these could be prevented. Case reviews and root cause analysis using a structured tool may help health care providers to learn from adverse events and to identify trends and recurring systems issues. We sought to establish the reliability of a root cause analysis computer application called Standardized Clinical Outcome Review (SCOR).

Methods: We designed a mixed methods study to evaluate the effectiveness of the tool. We conducted qualitative content analysis of five charts reviewed by both the traditional obstetric quality assurance methods and the SCOR tool. We also determined inter-rater reliability by having four health care providers review the same five cases using the SCOR tool.

Results: The comparative qualitative review revealed that the traditional quality assurance case review process used inconsistent language and made serious, personalized recommendations for those involved in the case. In contrast, the SCOR review provided a consistent format for recommendations, a list of action points, and highlighted systems issues. The mean percentage agreement between the four reviewers for the five cases was 75%. The different health care providers completed data entry and assessment of the case in a similar way. Missing data from the chart and poor wording of questions were identified as issues affecting percentage agreement.

Conclusion: The SCOR tool provides a standardized, objective, obstetric-specific tool for root cause analysis that may improve identification of risk factors and dissemination of action plans to prevent future events.

Résumé

Objectif : Des événements indésirables se manifestent dans jusqu’à 10 % des cas obstétricaux et jusqu’à la moitié de ces événements sont évitables. Les analyses de cas et l’analyse des causes fondamentales au moyen d’un outil structuré pourraient aider les fournisseurs de soins à tirer des leçons des événements indésirables et à identifier les tendances et les problèmes systémiques récurrents. Nous avons cherché à établir la fiabilité d’un logiciel d’analyse des causes fondamentales connu sous le nom de Standardized Clinical Outcome Review (SCOR).

Méthodes : Nous avons conçu une étude faisant appel à des méthodes mixtes pour évaluer l’efficacité de l’outil. Nous avons mené une analyse qualitative du contenu de cinq dossiers ayant été analysés tant au moyen des méthodes traditionnelles d’assurance de la qualité qu’au moyen de...
Learning from Adverse Events in Obstetrics: Is a Standardized Computer Tool an Effective Strategy for Root Cause Analysis?

INTRODUCTION

An adverse event, defined as an unexpected incident directly associated with the care of the patient, or an incident that results in injury or death, occurs in up to 10% of obstetric cases, and up to half of these could be prevented. Seventy percent of adverse events have been traced to failures of teamwork and communication.

The labour and delivery environment is uniquely vulnerable to adverse events due to the presence of multiple health care providers from a variety of disciplines, the acuity of cases, and the unpredictable timing of events. Further, the management of a particular case often brings together individuals who have not previously worked together.

Minimizing adverse events during the antenatal, intrapartum, and postnatal periods and developing a culture of patient safety within obstetrics requires effective strategies for implementing and measuring culture change.

Comprehensive multi-component programs for improving patient safety in obstetrics created by hospitals in the United States and the United Kingdom have demonstrated a reduction in the number of adverse events and the costs of compensating liability cases. In Canada, the Managing Obstetrical Risk Efficiently program sought to provide a comprehensive patient safety and professional development program for hospitals.

A key component of these comprehensive patient safety programs was the formal review of adverse events. MORE specifically recommended the use of peer case reviews and root cause analysis but did not provide specific details of how to conduct case reviews. The crux of formal RCA is to improve the identification of adverse events and to make the dissemination of lessons learned from the case more effective. RCA aims to determine what happened, why it happened, and to prevent future similar incidents. The central principle is that effective peer review is essential in improving practice. Standardized mechanisms for both identification of the cases requiring review and for conducting the review to identify risk factors and recommendations for action are critical to this process. There is evidence that systematic formal case reviews have a positive impact on the patient safety culture at an institution and on decreasing the rates of adverse events. A systematic review of interventions aimed at behaviour change within obstetrical practice demonstrated a positive impact of “audit and feedback” techniques in changing practice. In addition to the benefits of audit and feedback, there was evidence that standardizing the process itself was often a factor in the improved outcomes. Further, the use of a structured tool to investigate and learn from adverse events through RCA was recommended. Formal, standardized mechanisms for reporting adverse events and near misses facilitated the recognition of trends and addressed the failure to learn from critical incidents.

Despite the evidence of the impact of RCA and systematic case review, the current standard at most obstetric units in Canada involves the review of individual cases in isolation from larger systems issues and without a standardized approach. At our tertiary obstetric unit, which completed the MORE program over five years ago, it was the remit of the Obstetric Quality Assurance Committee to review cases involving adverse outcomes or near misses. The committee was interprofessional in its composition and included midwives, nurses, obstetricians, and paediatricians. The group met monthly or at the call of the chair. Cases requiring review were identified through an informal ad hoc process. The committee reviewed all maternal deaths and all unexpected stillbirths, together with any case brought to the attention of the chair. One member of the committee was assigned to review the clinical chart and present the case to the

Abbreviations

MORE Managing Obstetrical Risk Efficiently
OBSCQA Obstetric Quality Assurance
RCA root cause analysis
SCOR Standardized Clinical Outcome Review
UK United Kingdom

the group, and the case was then discussed by the committee, which made recommendations. Unfortunately, there was no method in place for systematically identifying all eligible records; there was no ability to track patterns or themes among similar cases from a systems perspective; and the accountability to prevent future events was not formalized. We recognized the shortcomings of this process, and piloted a new computer tool for RCA and systematic review, the Standardized Clinical Outcome Review.

The SCOR computer tool was developed in the UK following the review of several perinatal confidential enquiries that demonstrated significant variation in the ways in which National Health Service hospitals reviewed stillbirths and neonatal deaths. The review found that many poor outcomes were potentially avoidable, but the ad hoc approach to review of cases did not promote proper identification of key issues and clear learning or action points.16 A regional interprofessional working group was created to develop a standardized review and reporting mechanism, which would facilitate effective and efficient response to adverse outcomes. The aim was to use the application to provide a standardized process for RCA of perinatal deaths, to promote learning and action points that would improve practice, and to ensure that action points were implemented in a timely way. Additionally, when used by a region to track cases from more than one hospital, it facilitated the pooling of aggregate data to form a database to examine larger systematic trends.15 SCOR was launched in the UK in September 2011, and to date over 400 perinatal mortality cases have been entered and reviewed using the electronic tool.

The SCOR tool incorporates three key components:

1. systematic entry and assessment of data related to all phases of perinatal care through case review, including links to evidence-based guidelines to understand what happened;
2. automatic computer-generation of “case summary,” “key points,” “risk factors,” and “care issues” summarizing key factors that contributed to why an event happened (Figure 1); and
3. discussion and completion of an “action plan” by an interprofessional team, with the intention of preventing future events by outlining how care issues will be addressed and specifying a clear timeline and person responsible.

In September 2012, we began to work with our partners in the United Kingdom to create a Canadian version of the tool using Canadian research, hospital policies, and Society of Obstetricians and Gynaecologists of Canada

![Figure 1. Image of SCOR tool “risk factor” summary](image)
clinical practice guidelines that could be used for review of all adverse outcomes in obstetrics, as opposed to only perinatal deaths, as had been used in the UK.

We were interested in answering the following question: does the use of the SCOR computer application provide a feasible, acceptable, and reliable strategy for improving patient safety for mothers and newborns at our obstetric unit? We describe here the process of establishing the reliability of an adverse event RCA process through measures of inter-rater reliability and a comparative review of the computer process versus the traditional OBSDA.

**METHODS**

We implemented and evaluated a modified Canadian version of the SCOR tool between January 2013 and March 2014. A summary of the changes to the tool is provided in Table 1. We used a mixed method approach to answer our research question.

The feasibility and acceptability components of our research question were assessed following implementation with an interprofessional team which reviewed new cases involving adverse outcomes at our hospital between May 2013 and March 2014. This analysis is ongoing.

The effectiveness of the tool was evaluated in two ways (Figure 2). First, we carried out qualitative content and textual analysis based on retrospective comparative review of five adverse event cases that had previously been reviewed using the traditional OBSDA committee process.

The five cases occurred in 2009 and were selected by the chair of the OBSDA committee. First, the principal investigator of this study, who is not a member of the OBSDA group, read the full patient charts and entered the data for each chart into the SCOR computer application; a list of risk factors and care issues arising in each of the five cases was generated by the SCOR application. Next, this SCOR-generated output list was read line by line by the principal investigator and coded for keywords and recurring themes. The minutes and written documentation by the chair of the OBSDA committee, recorded at the time of the original review of the same charts, were then also read and coded line by line. The key words and recurring issues arising from the two documents were compared and contrasted. The codes were related and linked into meaningful clusters of similar concepts, which were then grouped together to form themes.

The second component of our data analysis was the determination of inter-rater reliability using the responses generated by four maternity care experts using SCOR for their review of five patient cases. Reviewer 1 (R1) and Reviewer 2 (R2) were members of the original OBSDA group that had reviewed the case previously. Both had been involved in obstetrics for over 25 years. Reviewer 3 (R3) and Reviewer 4 (R4), whom we describe as “blinded,” had not previously seen the chart and were not aware of any details of the clinical case. Both were experienced clinicians and had worked in obstetrics for over 10 years. Each reviewer, working independently, read the clinical chart for each case and entered his or her reviewer data into the SCOR application.

The SCOR tool contains 170 data entry points for the reviewer to complete during the standardized review. There is a combination of “factual” data entry questions (n = 52), such as maternal date of birth or infant birth weight, and questions that require clinical judgment or interpretation (n = 118), such as “Was the management of intrapartum risk factors appropriate?” The latter type of question has a drop-down menu of options for the reviewer to select an answer from, such as: “yes, plan was appropriate”; “yes, plan appropriate but not followed”; “no, plan not appropriate”; and “plan not identified.” The responses of R1 were compared with those of R2 by determining whether or not the responses to each question were identical. The same process was repeated for the responses of R3 and R4. Because of the categorical nature of the data set, we could not use traditional calculations such as Kappa to determine the inter-rater-reliability. Instead, we considered the mean percentage agreement rates for both the interpretive questions and the factual questions among the blinded and unblinded reviewers.

Ethics approval for the study was provided by the Hamilton Integrated Research Ethics Board.

Table 1. Summary of changes to SCOR tool

<table>
<thead>
<tr>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Expansion of scope of review from only perinatal deaths to include all adverse obstetric outcomes</td>
</tr>
<tr>
<td>• Addition of new “trigger” list of case identification to include all adverse obstetric events—maternal admission to ICU, maternal blood transfusion, maternal death, maternal return to OR/ L&amp;D, neonatal birth trauma, other fetal/neonatal morbidity, other maternal morbidity, termination of pregnancy</td>
</tr>
<tr>
<td>• Updated ‘Information Icons’ that reference best practice or current evidence to Canadian Clinical Practice Guidelines, local policies, or protocols</td>
</tr>
<tr>
<td>• Revised Canadian terminology, populations, measurements, and health care professionals</td>
</tr>
<tr>
<td>• Added questions and algorithm pathways for care outcomes other than perinatal deaths—including maternal recovery during postpartum, or neonatal concerns</td>
</tr>
</tbody>
</table>
Five adverse event cases were selected by the chair of the OBSQA committee for the retrospective comparative review and for inter-rater reliability. The cases occurred in 2009. One case involved an unexpected neonatal death, two cases involved management of atypical or abnormal fetal heart rate tracings, and two involved concerns about the timeliness of intrapartum management decisions and communication issues on the labour unit.

Qualitative Comparative Review
From the five charts, we compared the number and type of recommendations generated by the OBSQA committee and by the SCOR tool through qualitative content and textual analysis. Between the five cases, SCOR identified 107 risk factors and care issues, while the QA committee identified 25. When the retrospective comparative review was completed using qualitative analysis, three central themes that captured the differences between the SCOR standardized chart review and the traditional OBSQA review of the same five patient charts were identified.

The first theme identified during our comparative review related to consistency of documentation. We found that the format of documentation by the OBSQA committee was inconsistent and used a different format and even different terms for each of the five cases. The documentation included “notes,” “recommendations,” and “issues,” but these terms were not used consistently or with any rationale. In contrast, the SCOR application provided the same format of computer-generated “risk factors” and “care issues” for each case and provided a printed summary of the key facts.

The second theme pertained to the difference between facts and narrative descriptions. The notes from the OBSQA group often contained rich details and description of what happened in the case. However, these descriptions occasionally involved subjective details that may not have been known from reading the chart but that were provided by members of the interprofessional committee who may have been present or who had heard about the incident. Conversely, SCOR did not allow for a rich description of what happened; the data fields requested by the program only drew from what was written in the chart.

The final theme identified was the nature of the recommendations and action points arising from the chart review. The OBSQA committee focused on individual follow-up with each of the health professionals involved in the case. The recommendations often carried serious implications for the health professional involved, such as a change to their role or privileges at the institution. The SCOR recommendations were less focused on individual actions and highlighted small and large system issues. The large system issues addressed issues of staffing, education, and communication. Some of the small issues identified by SCOR were of little clinical relevance or were not linked to the outcome. For example, SCOR might identify as a risk factor that the woman was a smoker, but this risk did not contribute to the particular outcomes of the clinical case.

Inter-Rater Reliability
For the analysis of inter-rater reliability, we calculated percentage agreement between the two blinded reviewers and between the two unblinded reviewers (Table 2) who each answered 170 questions on the same five cases. Of the total 170 SCOR questions answered, we further
analyzed separately the 118 questions that required clinical interpretation or judgment and the 52 factual questions. Of the 3400 responses from the four reviewers on five cases and 170 questions, 1837 responses (54%) that were originally “not applicable,” “unknown,” or left blank were re-coded as a single response. These three responses were used interchangeably by reviewers when the relevant data could not be found in the chart. The relevance of each question and how it was answered depended on the case and the interpretation of the reviewer, who might have viewed an absence of information in the patient chart as either “not applicable” or “unknown.”

The mean percentage agreement rate for the five cases was similar between the two unblinded reviewers (76%, range 72% to 81%) and the two blinded reviewers (74%, range 65% to 81%) for 170 questions, suggesting that familiarity with the case did not necessarily affect how reviewers answered SCOR questions. The overall mean percentage agreement between all reviewers was 75% (range 65% to 81%).

The mean percentage agreement between all reviewers decreased to 69% (range 56% to 78%) when only the 118 interpretive questions were considered. In comparison, there was 89% (range 80% to 98%) agreement between all reviewers in the 52 factual questions.

**DISCUSSION**

Obstetric quality assurance and patient safety in obstetrics is a growing concern. Although adverse outcomes are rare, near misses and systems issues occur in up to 26% of births.13 Many poor outcomes are potentially avoidable, but the current process of RCA of near miss or adverse events does not ensure proper identification of either key issues or learning and action points for clinicians.16

Our pilot of the SCOR tool is one of the first standardized computer applications designed for review of adverse events in obstetrics in Canada. Findings from other disciplines have demonstrated that rates of adverse events decreased after formal RCA interventions.12 A systematic review by Percarpio et al. found that patient safety was improved following RCA, and that institutions that conducted less than four RCA reviews per year had higher rates of adverse events.17,18

Our findings from the retrospective comparative review and the inter-rater reliability indicate that the SCOR tool holds promise as a formal strategy for RCA. The strengths of the SCOR tool were that it identified a more complete and detailed picture of the risk factors and care issues that
may have contributed to why an event occurred, free from the subjectivity that often characterizes review of complex cases. This was demonstrated through the higher number of risk factors and action points generated per case. The tool allowed for clear tracking and definitions of action plans which are more likely to result in change. Another strength of the tool was that the SCOR recommendations and action plans were focused on systems level changes rather than being punitive or directing blame towards individuals. Shifting away from a culture of blaming and shaming is a key recommendation by many experts in patient safety. There is also evidence from a Cochrane review that patient safety interventions that are non-punitive and that include safe feedback for individuals involved in adverse outcomes were more likely to result in increased reporting of incidents and near misses. Health care professionals need to believe that the system of reporting and reviewing adverse events is safe and likely to result in quality improvements, or they will not report incidents.

One shortcoming of the reports generated through SCOR was that they lacked the richness of context and specific detail in describing what happened during an adverse event. The tool could be improved with the addition of more open-text entries in which reviewers can capture the story of how events unfolded. This may be particularly useful when a multidisciplinary team meets to discuss the case.

The percentage agreement between reviewers demonstrated that different health care providers completed the data entry and assessment of the case in a similar manner. This is an important finding, because the identification by the SCOR of risk factors, care issues, and action plans was based upon how the individual reviewing the case answered both the interpretive and factual questions. This validates the conclusion that, for the most part, the tool asks appropriate questions and provides effective and clear responses. The testing process highlighted individual questions that had poor user agreement. In most cases, disagreement was due to a misinterpretation of either the question or the possible answers from the pull-down menu. Identifying questions with high individual disagreement rates will permit exploration of the utility and sensitivity of the question and potential re-wording or modification. Our evaluation also identified questions which needed to be removed from the tool because of a lack of clarity or an emphasis on practices or documentation that do not have a Canadian equivalent. For example, the tool contained a question that asked if the mother was born outside Canada. The documentation of this fact is up to the discretion of the clinician on the Ontario Antenatal Record, whereas in the UK it is required to ask if the mother was born outside the country. Reviewers in our pilot were inconsistent in their completion of such questions, which affected the percentage agreement.

Structured RCA of cases through a standardized tool has the ability to reduce system errors and to improve the culture of patient safety through audit, feedback, and dissemination of lessons learned. Researchers in the UK found that the quality of chart audit improved following the implementation of standardized tools, and that patient satisfaction with care also subsequently improved. From discussion with our colleagues in the UK, we learned that standardized reviews of adverse perinatal outcomes helped to identify underlying preventable causes, and they have underpinned significant reductions in stillbirths in the West Midlands of England. Recent evaluation of a pilot of the SCOR application in 17 maternity units in the United Kingdom found it to be a useful tool for assessing and learning lessons from adverse outcomes.

One of the challenges of testing this tool was that there was no accepted standard for comparison. The traditional quality assurance process of chart audit and feedback by an obstetrical committee is also untested, and it would be difficult to know what percentage agreement would exist among reviewers using that method. For this reason, we also conducted questionnaire and focus group data with members of the interdisciplinary team who used SCOR instead of the traditional OBSQA process to understand the acceptability and feasibility of the use of the tool. This analysis is forthcoming.

One of the limitations of our study was the small number of cases reviewed. In addition, because of the nature of the categorical SCOR questions it was difficult to use traditional tests of reliability. Following revision of the questions based on the work of this pilot, additional testing of the tool with more cases and more reviewers would be beneficial. Further research on SCOR should focus on the number and type of action points generated by the tool and the number of action points completed, to ensure that it provides a useful strategy for tracking and responding to events. Another logical next step for evaluating the tool would be to measure the impact of the tool on the rates of adverse events, rates of incident reporting, and the culture of patient safety at a labour and delivery unit before and after implementation.

CONCLUSION

The SCOR computer application provides a standardized, obstetric-specific tool for RCA that may help promote a culture of patient safety through improved identification of
Learning from Adverse Events in Obstetrics: Is a Standardized Computer Tool an Effective Strategy for Root Cause Analysis?

the risk factors and care issues that lead to adverse events, and through improved dissemination of lessons learned using clear and specific action plans. Although the tool requires additional modification for the Canadian setting, it appears to be a rigorous and reliable method for standardizing the process of patient case review and RCA. Implementation in more labour and delivery units would be a useful next step for determining the impact of the tool on the number of adverse events and the wider culture of patient safety. Further research will focus on comparing the SCOR-generated risk factors and recommendations for a case that arise from the entries made by two different reviewers, in order to determine the application’s sensitivity to differences in interpretation of the case by different clinicians.

REFERENCES