

AHRQ Safety Program for Perinatal Care

Summary Report

Prepared for:

Agency for Healthcare Research and
Quality, U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20850
www.ahrq.gov

Contract No. HHS290201000024I

Prepared by:

Research Triangle International
Research Triangle Park, NC

AHRQ Publication No. 17-0003-24-EF

May 2017



Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov



This project was funded under contract number HHS290201000024I from the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services.

The opinions expressed in this document are those of the authors and do not reflect the official position of AHRQ or the U.S. Department of Health and Human Services. None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this document.

This document may be used and reprinted without permission except those copyrighted materials that are clearly noted in the document. Further reproduction of those copyrighted materials is prohibited without the express permission of copyright holders.

Suggested citation: Sorensen, A.V., Webb, J., Clare, H.M., Banger, A., Jacobs, S., McArdle, J., Pleasants, E., Burson, K., Lasater, B., Gray, K., Poehlman, J., Kahwati, L.C., AHRQ Safety Program for Perinatal Care: Summary Report. Prepared under Contract No. 290201000024I (RTI International). AHRQ Publication No. 17-0003-24-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2017. www.ahrq.gov/perinatalsafety.

Contents

1.	Introduction	1
2.	Program Description and Implementation	3
2.1	Program Design and Customizable Toolkit.....	3
2.2	Unit Recruitment and Retention	5
2.3	Implementation Support	6
3.	Evaluation Methods	8
3.1	Overview.....	8
3.2	Implementation Evaluation Methods	9
3.2.1	Quantitative Data Collection, Measures, and Analysis	9
3.2.2	Qualitative Data Collection and Analysis.....	10
3.2.3	Mixed Methods Analyses.....	10
3.3	Impact Evaluation Methods	11
3.3.1	Quantitative Data Collection and Analysis	11
3.3.2	Qualitative Data Collection and Analysis.....	12
4.	Evaluation Findings	13
4.1	Implementation Evaluation Findings	13
4.1.1	Teamwork and Communication.....	13
4.1.2	Perinatal Safety Strategies	14
4.1.3	In Situ Simulations	16
4.1.4	Overall Implementation Effectiveness	17
4.2	Impact Evaluation Findings	18
4.2.1	Impact on Unit Patient Safety Culture	18
4.2.2	Impact on Adverse Events.....	19
5.	Discussion	22
5.1	Summary of Findings	22
5.1.1	Overall Program Implementation.....	22
5.1.2	Teamwork and Communication.....	23
5.1.3	Perinatal Safety Strategies	23
5.1.4	In Situ Simulations	24
5.1.5	Limitations of the Implementation Evaluation.....	24
5.1.6	Summary of Impact on Unit Patient Safety Culture.....	25
5.1.7	Summary of Impact on Adverse Events	25
5.1.8	Limitations of the Impact Evaluation.....	25
5.2	Recommendations for Future Program Design, Implementation, and Evaluation	27
5.2.1	Program Design.....	27
5.2.2	Program Implementation	27
5.2.3	Program Evaluation	28
6.	References	30
	List of Acronyms/Abbreviations.....	32

Figures

Figure 1. AHRQ SPPC implementation and evaluation logic model, 2015–2016.....	2
Figure 2. Contents of the customizable program toolkit for the AHRQ SPPC, 2015–2016	4
Figure 3. Recruitment and participation of L&D units in the AHRQ SPPC, 2015–2016	5
Figure 4. Contextual implementation features evaluated in the AHRQ SPPC, 2015–2016	11
Figure 5. Maternal and neonatal adverse events for impact evaluation in the AHRQ SPPC, 2015–2016	12
Figure 6. Implementation of key CUSP principles, strategy-specific teamwork and communication, and in situ simulations for selected perinatal safety strategies, AHRQ SPPC, 2015–2016	16

Tables

Table 1. Characteristics of the 46 L&D units that completed participation in the AHRQ SPPC, 2015–2016	6
Table 2. Summary of data and data sources used for the evaluation of the AHRQ SPPC, 2015–2016	8
Table 3. Summary of impact on maternal and neonatal adverse events for L&D units submitting adverse event data at baseline and followup (N=43) for the AHRQ SPPC, 2015–2016	20

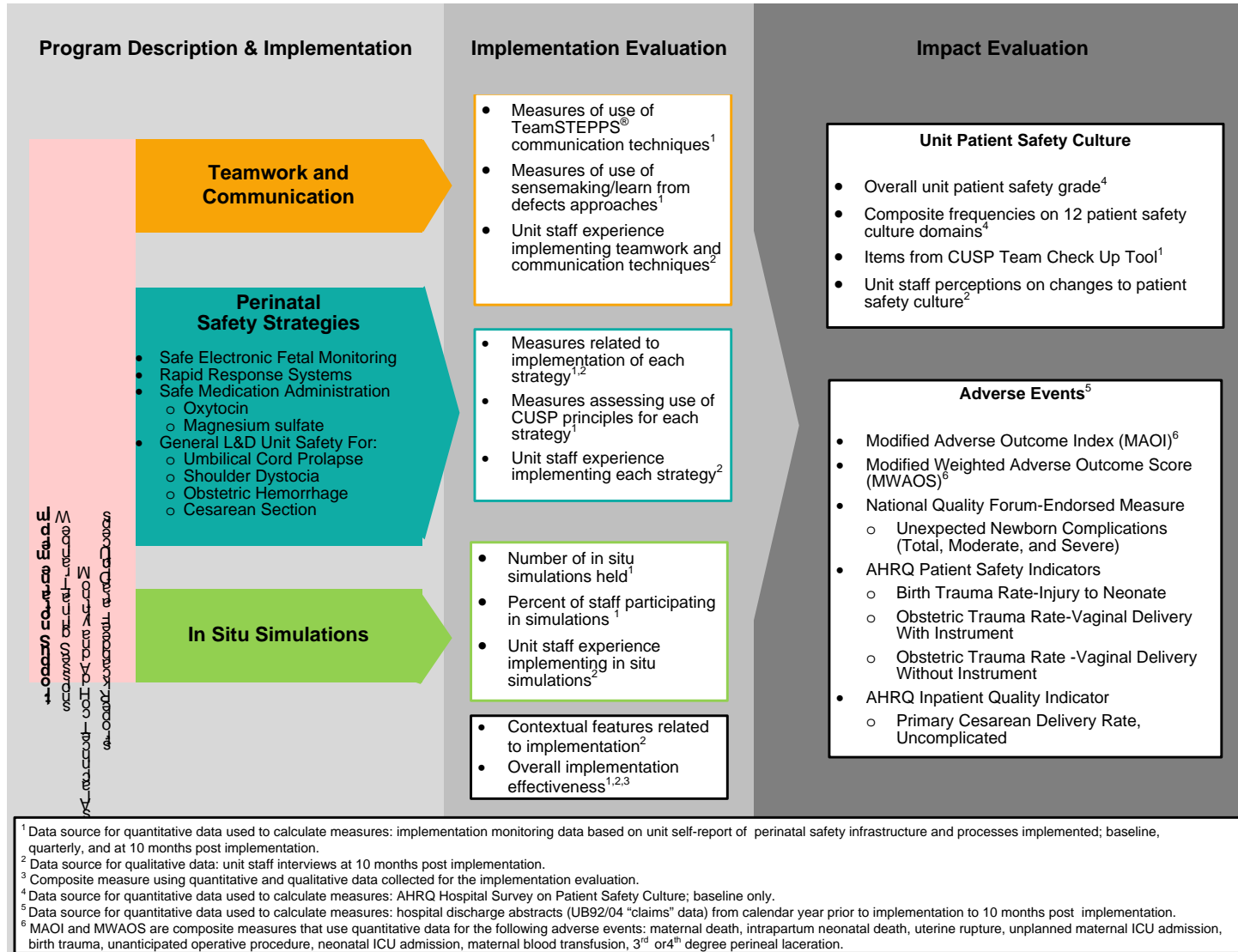
1. Introduction

Of the 3.9 million births in the United States each year,¹ 2 percent are estimated to involve an adverse event; at least half are potentially preventable.² A review by the Joint Commission found that, between 2004 and 2014, poor communication was a root cause of 48 percent of sentinel maternal events and 70 percent of sentinel neonatal events.³ In addition to communication failures, patients on labor and delivery (L&D) units are at risk of medication errors due to the frequent use of high-alert medications, and though obstetric emergencies are rare, they have the potential to result in catastrophic outcomes such as maternal or neonatal death if an appropriate clinical response is not provided in a safe, coordinated, and timely manner. Lastly, the use of inappropriate interventions, or interventions provided in an unsafe manner, also increases the risk of adverse events on L&D units. High-reliability systems and a culture of learning from errors (or near misses) are needed to minimize preventable harms.

The Agency for Healthcare Research and Quality (AHRQ) developed the Safety Program for Perinatal Care (SPPC) in order to improve the patient safety culture of L&D units and decrease maternal and neonatal adverse events resulting from poor communication and system failures. This program extends AHRQ's existing Comprehensive Unit-based Safety Program to L&D units and offers hospitals an approach to improving the unit patient safety culture and obstetric care processes to eliminate patient safety failures as a cause for adverse maternal and neonatal events. Program design, implementation support, and evaluation were provided by a national team composed of staff from AHRQ, RTI International, and the National Perinatal Information Center in addition to expert clinical faculty from a variety of organizations.

The logic model depicted in **Figure 1** guided program design, implementation, and evaluation. The program design and implementation included three pillars: teamwork and communication skills, selected perinatal safety strategies, and in situ simulation training. A toolkit was developed to support L&D unit implementation of these program pillars. Implementation support (e.g., training, technical assistance, data feedback reports) was provided to 46 L&D units across 10 States that participated in a nationally coordinated implementation of the program. A mixed methods evaluation was also conducted to examine the implementation of the program, including activities in each of the three pillars, and the impact of the program on unit patient safety culture and maternal and neonatal adverse events. This Summary Report provides an overview of the program design and implementation (**Section 2**), the methods used for evaluation (**Section 3**), the evaluation findings (**Section 4**), and discussion, including recommendations for future perinatal safety programs (**Section 5**).

Figure 1. AHRQ SPPC implementation and evaluation logic model, 2015–2016



Abbreviations: AHRQ = Agency for Healthcare Research and Quality, CUSP = Comprehensive Unit-based Safety Program, ICU = intensive care unit, L&D = Labor and Delivery, MAOI = modified Adverse Outcome Index, MWAOS = Modified Weighted Adverse Outcome Score

2. Program Description and Implementation

The Agency for Healthcare Research and Quality (AHRQ) Safety Program for Perinatal Care (SPPC) involved three distinct phases: design of the program and development of the customizable program toolkit, recruitment of labor and delivery (L&D) units to participate in a nationally coordinated implementation, and implementation support provided to participating units. The following sections expand upon each of these phases.

2.1 Program Design and Customizable Toolkit

The AHRQ SPPC was based on the three pillars shown in **Figure 1**. These pillars represent mutually reinforcing and interdependent activities that L&D units can use to improve perinatal safety on their units. A customizable toolkit was developed to support the implementation of these three program pillars. Of note, the toolkit was not designed to recommend clinical practice or develop clinical guidelines for obstetric conditions; rather, it was designed to facilitate improved quality and patient safety on L&D units through improved teamwork and communication, use of specific perinatal safety strategies, and reinforcement of teamwork and communication and safe obstetric care practices through in situ simulations.

The toolkit contents (**Figure 2**) were based on an evidence review, two technical expert panels, and testing during a pilot implementation phase. The toolkit was designed to be similar to other AHRQ Comprehensive Unit-based Safety Program (CUSP) initiatives in terms of structure, look, and feel. AHRQ's existing TeamSTEPPS[®] teamwork and communication platform serves as the foundation of the teamwork and communication pillar. Four perinatal safety strategies were offered: safe electronic fetal monitoring; rapid response systems; safe medication administration; and general L&D unit safety, which includes options that target specific obstetric conditions or procedures as indicated in **Figure 2**. Tools developed to support the implementation of the perinatal safety strategies pillar focused on demonstrating how to apply selected CUSP principles (i.e., standardizing, creating independent checks, learning from defects, and engaging patients and families) to these specific strategies. Tools developed for the in situ simulation pillar offered guidance to support the implementation of in situ simulation training, including sample simulation scenarios and a 30-minute documentary-style video.

Figure 2. Contents of the customizable program toolkit for the AHRQ SPPC, 2015–2016

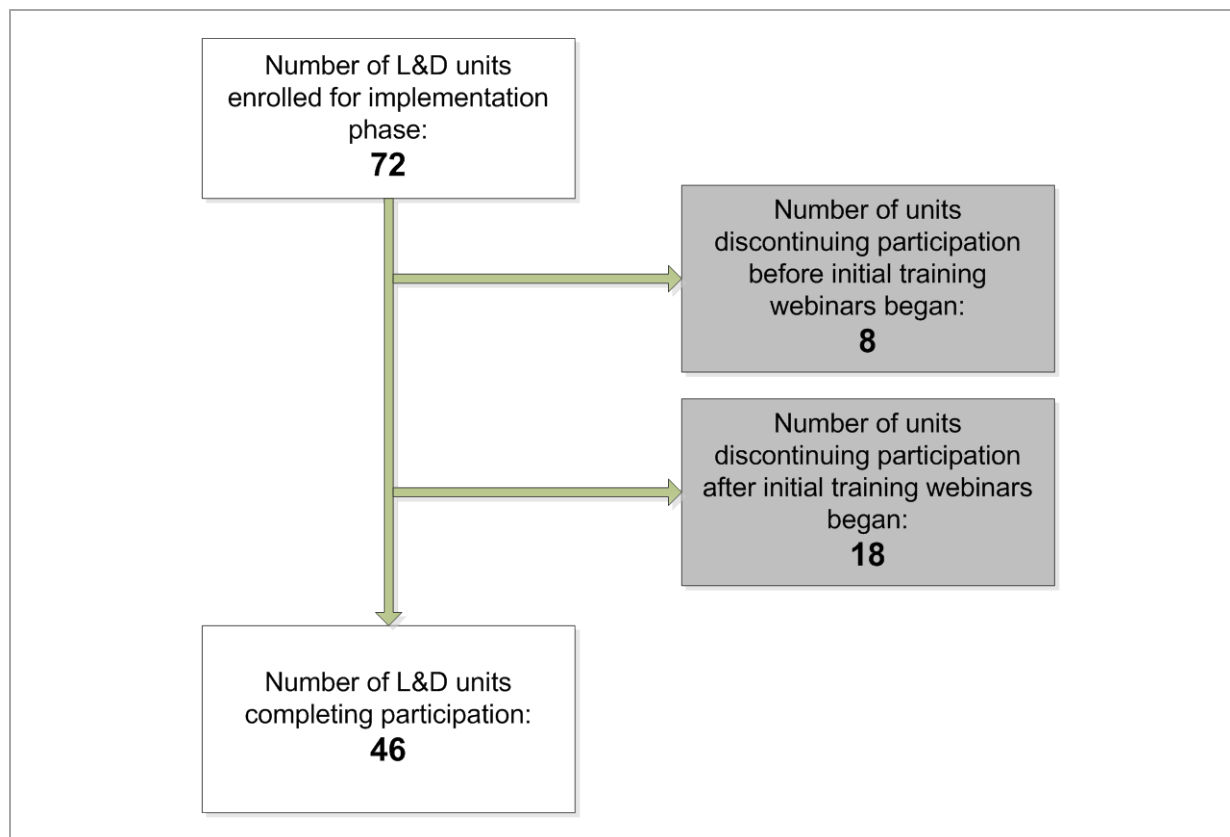
<p>Teamwork and Communication</p>	<p>Learn About the Comprehensive Unit–Based Safety Program (CUSP)</p> <ul style="list-style-type: none"> • Slide Presentation & Facilitator Notes <p>Assemble Team & Engage Leadership</p> <ul style="list-style-type: none"> • Slide Presentation & Facilitator Notes <p>Implement Teamwork & Communication</p> <ul style="list-style-type: none"> • Slide Presentation & Facilitator Notes <p>Understand the Science of Safety</p> <ul style="list-style-type: none"> • Slide Presentation & Facilitator Notes <p>Sensemaking & Learn From Defects</p> <ul style="list-style-type: none"> • Slide Presentation & Facilitator Notes <p>Engage Patients & Families</p> <ul style="list-style-type: none"> • Slide Presentation & Facilitator Notes <p>Supporting CUSP Tools</p> <ul style="list-style-type: none"> • Am I Ready To Become an Advisor • Background Quality Improvement Team Information Form • Be a Partner in Your Care • Board Checklist • CEO/Senior Leader Checklist • Culture Checkup Tool • Safety Issues Worksheet for Senior Executive Partnership • Shadowing Another Professional Tool • Staff Safety Assessment • Team Checkup Tool • Working With Patient & Family Advisors
<p>Perinatal Safety Strategies</p>	<p>Monitoring for Perinatal Safety</p> <p>Electronic Fetal Monitoring (EFM)</p> <ul style="list-style-type: none"> • Slide Presentation & Facilitator Notes • Tool: Safe EFM Pocket Card <p>Safe Medication Administration</p> <ul style="list-style-type: none"> • Slide Presentation & Facilitator Notes • Tool: Safe Oxytocin Administration • Tool: Safe Magnesium Sulfate Administration <p>Rapid Response for Perinatal Safety</p> <p>Obstetrical Rapid Response</p> <ul style="list-style-type: none"> • Slide Presentation & Facilitator Notes • Tool: Rapid Response Systems <p>L&D Unit Safety</p> <ul style="list-style-type: none"> • Slide Presentation & Facilitator Notes • Tool: Safe Cesarean Section • Tool: Obstetric Hemorrhage • Tool: Shoulder Dystocia • Tool: Umbilical Cord Prolapse
<p>In Situ Simulation</p>	<p>Getting Started</p> <ul style="list-style-type: none"> • Slide Presentation & Facilitator Notes • Tool: Facilitation Instructions for In Situ Simulation • Tool: Training Video <p>Simulation Scenarios</p> <ul style="list-style-type: none"> • Postpartum Hemorrhage • Shoulder Dystocia • Umbilical Cord Prolapse • Antepartum Hemorrhage • Preeclampsia/Seizure • Uterine Tachysystole • Magnesium Toxicity • Vaginal Birth After Cesarean (VBAC)/Abdominal Pain • Postoperative Cesarean Section Complication

Abbreviations: CEO = chief executive officer, CUSP = Comprehensive Unit-based Safety Program, EFM = electronic fetal monitoring, L&D = Labor and Delivery, VBAC = vaginal birth after cesarean

2.2 Unit Recruitment and Retention

The national team recruited L&D units from across the country to participate in the implementation of the program. Recruitment efforts targeted State hospital associations, regional perinatal quality and safety collaboratives, Centers for Medicare and Medicaid Services Hospital Engagement Networks, and health care systems. In turn, these coordinating entities facilitated recruitment of hospitals among their members. By December 2014, there were 72 L&D units located across 10 States and representing nine coordinating entities enrolled in the program. Over the course of implementation, 26 L&D units discontinued participation, resulting in 46 units that completed the full implementation phase. As illustrated in **Figure 3**, 8 units discontinued participation before initial training Webinars began; 18 units discontinued participation after these Webinars began.

Figure 3. Recruitment and participation of L&D units in the AHRQ SPPC, 2015–2016



Abbreviation: L&D = Labor and Delivery

The most common reasons for discontinuation were participation in competing national, State, or internal hospital quality improvement initiatives; inadequate unit staffing and staff turnover; high data burden pertaining to program evaluation activities; and fears of litigation resulting from sharing data with external entities for program evaluation. Characteristics of the 46 L&D units that completed participation are summarized in **Table 1**.

Table 1. Characteristics of the 46 L&D units that completed participation in the AHRQ SPPC, 2015–2016^a

Hospital or L&D Unit Characteristic	Frequency (%) or Mean (SD)
Publicly owned	9 (20%)
Rural referral center	5 (11%)
Mean number of hospital beds	321 (277)
Mean number of L&D unit beds	10 (6)
Level 1 Basic neonatal care	14 (31%)
Level 2 Specialty neonatal care	12 (27%)
Level 3 Subspecialty neonatal care ^b	18 (40%)
Graduate medical program in obstetrics and gynecology	12 (27%)
Annual number of births	2,077 (2,327)
Percent of births that are primary cesarean sections	17 (4)
Mean age of L&D unit patients	28 (2)
Percent of L&D patients that are nonwhite	34 (24)

^a Characteristics for units that did not complete participation were not available because these units did not submit any baseline data.

^b Includes Level 3A, 3B, 3C, and 3D. 2 L&D units did not report level of neonatal care.
Abbreviations: L&D = Labor and Delivery, SD = standard deviation

2.3 Implementation Support

The national team supported L&D units' implementation of the program through an initial series of Webinar trainings, monthly and ad hoc technical assistance Webinars and calls, and unit-specific data feedback reports.

Each unit formed a local implementation team to guide implementation activities. Members of the local implementation teams were invited to attend an initial series of five Webinar trainings. Each Webinar reviewed core CUSP principles as applied to the perinatal safety context, discussed practical approaches to implementation, and provided an overview of available tools and resources in the toolkit that L&D units could use to support their local implementation. After the initial series of Webinar trainings, L&D units were instructed to begin their implementation. First and foremost, units were instructed to develop or continue to build staff competency with TeamSTEPPS teamwork and communication skills. This was considered the foundational pillar of the program. Next, units were instructed to (1) select and implement at least two perinatal safety strategies and (2) reinforce teamwork and communication and the use of perinatal safety strategies through in situ simulations. Units were instructed to select the perinatal safety strategies that best met their local culture and needs.

Monthly technical assistance Webinars reinforced the content of the customizable toolkit and facilitated peer-to-peer learning. In addition, ad hoc calls with individual L&D units addressed specific challenges faced by those units. Support for implementation also included L&D unit staff access to a user support network, which was a Web-based repository for all toolkit resources and recorded Webinar sessions. To maintain engagement with L&D units throughout the implementation phase, a weekly email newsletter was disseminated to participating units that highlighted selected toolkit resources, offered tips for implementation and problem-solving barriers, and reminded units of deadlines for data reporting for

evaluation. The national team compiled implementation monitoring data submitted by each participating L&D unit (described in **Section 3**) into an individualized data feedback report to provide L&D units with a summary of process and clinical measures related to perinatal safety prior to, during, and after program implementation.

3. Evaluation Methods

3.1 Overview

The evaluation of the Agency for Healthcare Research and Quality (AHRQ) Safety Program for Perinatal Care (SPPC) examined both *implementation* and *impact*, as indicated in **Figure 1**. The implementation evaluation characterized the implementation process and experience, and the impact evaluation determined the effect of the program on the unit patient safety culture and maternal and neonatal adverse events. A mixed methods approach was used for both evaluations and relied on qualitative and quantitative data collection and analysis. A summary of the data and data sources used for the evaluation is provided in **Table 2**. Labor and Delivery (L&D) unit and hospital characteristics, such as level of neonatal care, number of annual births, type of hospital ownership, rural designation, characteristics of the patient population served by L&D unit, and the use of resident physicians on the unit, were collected at baseline for characterizing the participating L&D units and for stratifying findings in both the implementation and impact evaluations.

Table 2. Summary of data and data sources used for the evaluation of the AHRQ SPPC, 2015–2016

Data	Type of Data	Data Source	Implementation Evaluation			Impact Evaluation
			Teamwork and Communication	Perinatal Safety Strategies	In Situ Simulations	
Implementation of perinatal safety infrastructure and processes	Quantitative	Unit reporting on a Web-based form	✓	✓	✓	
Implementation experiences	Qualitative	Semi-structured interviews with participating units	✓	✓	✓	
		Structured interview debriefing form	✓	✓	✓	
Hospital and L&D characteristics	Quantitative	AHA Annual Member Survey; unit reporting on a Web-based form	✓	✓	✓	✓ ^a
Unit patient safety culture	Quantitative	AHRQ Hospital Survey on Patient Safety Culture				✓ ^b
	Quantitative	CUSP Team Checkup Tool on a Web-based form				✓
	Qualitative	Semi-structured interviews with participating units				✓
Adverse maternal and neonatal events	Quantitative	Hospital discharge abstracts (UB92/04 claims data)				✓

^a Hospital and L&D characteristics submitted at baseline were used to stratify the adverse event impact analyses.

^b Data from this survey were submitted only at baseline; thus, these data were only used to stratify adverse event impact analyses.

Abbreviations: AHA = American Hospital Association, AHRQ = Agency for Healthcare Research and Quality, CUSP = Comprehensive Unit-based Safety Program, L&D = Labor and Delivery

3.2 Implementation Evaluation Methods

The implementation evaluation involved quantitative, qualitative, and mixed methods data collection and analysis. As part of program participation, all L&D units were instructed to fully implement the teamwork and communication and in situ simulation pillars. L&D units were also instructed to implement at least two strategies from the perinatal safety strategy pillar. Thus, L&D units varied with respect to which strategies they implemented and how they chose to customize the implementation of each selected strategy.

3.2.1 Quantitative Data Collection, Measures, and Analysis

L&D units submitted quantitative data related to implementation processes for the national team to monitor and to use in generating unit-specific data feedback reports provided as part of implementation support (**Section 2.3**). L&D units submitted several types of quantitative data about the infrastructure and processes used for implementation of each pillar. These data were collected at four time points: baseline, the end of the first quarter after the start of implementation, the end of the second quarter, and 10 months after the start of implementation.

Units reported implementation monitoring data on a Web-based form; these data were used to calculate implementation measures for each pillar. Examples of measures used for each pillar are below:

- Teamwork and communication
 - proportion and types of staff trained in TeamSTEPPS®
 - use of specific TeamSTEPPS techniques (e.g., huddles, debriefs) by staff on the L&D unit
- Perinatal safety strategies
 - the extent to which Comprehensive Unit-based Safety Program (CUSP) principles (i.e., standardizing, creating independent checks, learning from defects) had been applied to the unit's selected perinatal safety strategies
 - Strategy-specific clinical process measures (e.g., proportion of cases of obstetric hemorrhage in which quantification of blood loss was used, and proportion of births with electronic fetal monitoring findings documented using standard nomenclature)
- In situ simulations
 - proportion and types of staff participating in in situ simulations
 - frequency of in situ simulations
 - monitored and multidisciplinary participation in in situ simulation
 - variety of simulation scenarios used

These pillar-specific implementation measures were aggregated across all units with available data for the four reporting periods. These measures were also stratified at baseline and at final based on several L&D unit or hospital characteristics: coordinating entity (e.g., State hospital associations); the annual number of births (three categories); the level of neonatal care (three categories)⁴; and the use of resident physicians on the L&D units.

3.2.2 *Qualitative Data Collection and Analysis*

The purpose of collecting qualitative data related to program implementation was to gain a more comprehensive and nuanced understanding of the implementation process than would be possible using only quantitative implementation monitoring data.

The main sources of qualitative data for the implementation evaluation were transcripts of interviews conducted in person or by phone with 131 L&D unit staff from 45 sites.^a Interview protocols were semi-structured and guided by constructs from the Consolidated Framework for Implementation Research (CFIR).⁵ The protocols included questions about the implementation experience for teamwork and communication, perinatal safety strategies, and in situ simulation pillars. Questions pertaining to perinatal safety strategies were tailored based on those selected by that unit. Each interview audio recording was transcribed and coded for emerging themes analyses and interpretation. Further, key interview findings were captured into a structured form immediately after the interview to summarize key contextual features related to implementation at each unit. In addition to interview transcripts, notes taken by the national team during Webinar training sessions and technical assistance calls were used in qualitative analyses.

3.2.3 *Mixed Methods Analyses*

Several mixed methods analyses contributed to the implementation evaluation in order to identify characteristics or features associated with effective implementation overall. Quantitative data from implementation monitoring were combined with qualitative data from unit interviews to create a composite measure of implementation effectiveness for each unit. The purpose of this measure was to have a single measure to characterize implementation of the overall program. For a unit to be characterized as having effective implementation, it needed evidence from the quantitative and qualitative data that each of the three program pillars had been implemented. Units without such evidence were considered as not having effective implementation. This composite measure of implementation effectiveness was used as the outcome in a qualitative comparative analysis (QCA) of the relationship between several contextual features derived from CFIR (**Figure 4**) and effective implementation. Qualitative comparative analysis is nonstatistical technique that can combine qualitative and quantitative data within an analysis to identify complex causal relationships among variables and an outcome. The composite measure of effective implementation was also used as the outcome in several multivariate quantitative analyses assessing the independent association of implementation effectiveness and these contextual features, and with other hospital or L&D unit characteristics, such as annual number of births, level of neonatal care, baseline patient safety culture, and presence of other perinatal quality or safety initiatives at baseline.

^a Two of the 46 units that completed 10 months of program participation were not interviewed because of staff turnover on 1 unit and lack of available unit staff time to participate in the interview at the other unit. One of the 26 sites that did not complete 10 months of program participation was interviewed to provide additional information regarding implementation challenges. Thus, interviews were conducted with 45 units in total.

Figure 4. Contextual implementation features evaluated in the AHRQ SPPC, 2015–2016

- High engagement of hospital leadership in implementation
- L&D unit adequately resourced to implement program
- Dedicated and adequate time for unit leaders to implement program
- Appropriate kinds of unit staff engaged in implementation
- Assistance/resources external to the hospital received in support of implementation

Abbreviation: L&D = Labor and Delivery

3.3 Impact Evaluation Methods

The impact evaluation involved quantitative and qualitative data collection and analysis designed to evaluate the broad (i.e., not pillar-specific) impact of the program on the units' patient safety culture and on the frequency of maternal and neonatal adverse events.

3.3.1 Quantitative Data Collection and Analysis

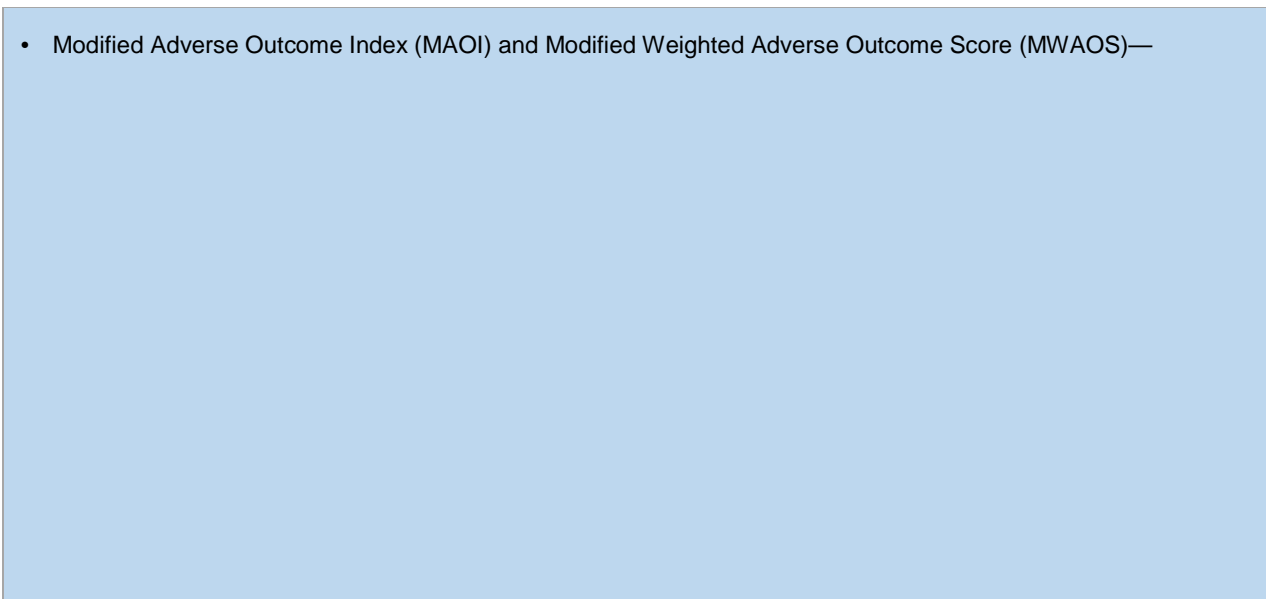
L&D units submitted several types of quantitative data for use in the impact evaluation. These data were collected at four time points: baseline, the end of the first quarter after the start of implementation, the end of the second quarter, and at 10 months after the start of implementation, though not all impact data were collected at each of these time points. Impact measures were calculated based on data that were aggregated across all units with available data. These measures were also stratified based on several L&D unit or hospital characteristics; these included coordinating entity (e.g., State hospital associations), the annual number of births (three categories), the level of neonatal care (three categories)⁴, and the use of resident physicians on the L&D units (yes or no).

Impact on unit patient safety culture. The main quantitative data source for evaluating unit patient safety culture was the CUSP Team Checkup Tool, which units submitted on a Web-based form. This tool, which was reproduced directly from the AHRQ CUSP Toolkit, includes 18 individual items across three domains: knowledge/skills, attitudes/beliefs, and resources. In addition to this tool, units fielded and submitted data from the AHRQ Hospital Survey on Patient Safety Culture at baseline. This survey, which consists of 43 individual items across 12 domains, was made optional for submission at 10 months to reduce data burden on participation units. Thus, data from the baseline safety culture survey were only used to stratify the adverse event impact analyses, since followup culture survey data was submitted by only 2 of the 46 units that completed participation.

Impact on adverse events. The data source for evaluating impact on adverse events was hospital discharge abstracts (i.e., claims data from the UB92/04) that units submitted for the calendar year prior to implementation and for the 10-month period following the start of implementation. The adverse events that were evaluated using these data are summarized in **Figure 5** and include two composite measures of maternal and neonatal adverse events, two maternal adverse event measures, two neonatal adverse event measures, and one inpatient quality reporting measure. Absolute and relative changes in these events from the calendar year prior to implementation to 10 months

postimplementation were calculated. Changes in the absolute incidence of adverse events were determined to be statistically significant when the P value was less than 0.05 using the Wilcoxon signed-rank test for non-normally distributed data. These measures were also stratified based on several L&D unit or hospital characteristics; these included coordinating entity (e.g., State hospital associations), the annual number of births (three categories), the level of neonatal care (three categories)⁴, and the use of resident physicians on the L&D units (yes or no).

Figure 5. Maternal and neonatal adverse events for impact evaluation in the AHRQ SPPC, 2015–2016



- Modified Adverse Outcome Index (MAOI) and Modified Weighted Adverse Outcome Score (MWAOS)—

Abbreviations: ICU = intensive care unit, MAOI = modified Adverse Outcome Index, MWAOS = Modified Weighted Adverse Outcome Score, NICU = neonatal intensive care unit

3.3.2 *Qualitative Data Collection and Analysis*

The purpose of collecting qualitative data related to program impact was to gain a more in-depth and nuanced understanding of the impact of the program on the units' patient safety culture than is possible using only quantitative data. The main sources of qualitative data were transcripts of interviews conducted in person or by phone with 131 L&D unit staff from 45 sites, as previously described in **Section 3.2.2**. Each interview was coded for analyses of emerging themes related to the overall impact of the program on the units' patient safety culture.

4. Evaluation Findings

This section describes findings from the *implementation* evaluation and the *impact* evaluation. The implementation evaluation describes findings about the units' local experiences with program implementation. These findings are based on quantitative data collected from units on perinatal safety infrastructure and processes and on qualitative data collected from interviews with participating Labor and Delivery (L&D) unit staff at the end of the 10-month implementation. These findings are organized by the three program pillars. The impact evaluation section describes the impact of the program overall and is organized by impact on the units' patient safety culture and impact on maternal and neonatal adverse events.

4.1 Implementation Evaluation Findings

4.1.1 *Teamwork and Communication*

Interpreting trends in the proportion of staff trained with TeamSTEPPS® from the implementation monitoring data was challenging as many units did not submit these data at baseline. During the interviews at the end of 10-month implementation, most units reported that they trained multiple unit staff in TeamSTEPPS but faced significant challenges tracking and reporting numbers of staff trained as part of implementation monitoring.

Many units reported adopting several TeamSTEPPS communication techniques. Units used different pathways to improve teamwork and communication, and different approaches to using TeamSTEPPS techniques. Many used the traditional train-the-trainer approach and relied on TeamSTEPPS master trainers. Units without master trainers used a condensed 2-hour training developed by the national implementation team as an interim strategy until staff at the unit could attend TeamSTEPPS master training. TeamSTEPPS implementation within units was typically driven by local nurse facilitators, though some L&D units drove adoption through mandated training. Once local facilitators completed initial TeamSTEPPS trainings for unit staff, they sustained the use of TeamSTEPPS tools and techniques through peer support and positive peer pressure, leading by example, and by deploying multiple checks and reminders. The implementation of this pillar allowed L&D units to either introduce these teamwork concepts or “resurface” them in situations where prior implementation efforts had not led to sustainment of their use. The most commonly adopted TeamSTEPPS techniques were SBAR (situation, background, assessment, recommendation), huddles, and debriefs. Most units planned to continue training additional staff in TeamSTEPPS or introduce additional TeamSTEPPS techniques. A number of units established a process for sustaining the use of this new skillset by embedding it into an annual or semiannual training requirement, and by requiring all new staff to receive TeamSTEPPS training. A few units even formalized some of the TeamSTEPPS techniques into documentation

templates within their electronic health record systems, thereby reinforcing their use and ensuring sustainability.

4.1.2 *Perinatal Safety Strategies*

This section describes unit experiences with implementation of perinatal safety strategies that focus on common care processes or procedures provided on L&D units (e.g., cesarean section, oxytocin administration, electronic fetal monitoring (EFM)) and on less common but urgent obstetric conditions requiring a rapid clinical response (e.g., deliveries complicated by umbilical cord prolapse, shoulder dystocia, or obstetric hemorrhage). As described during interviews following the 10-month implementation, units used diverse approaches targeting changes to unit policies, processes, and documentation. Changes for some strategies (e.g., safe medication administration and rapid response systems) required new equipment and collaboration with other units (e.g., pharmacy, emergency department). For other strategies (e.g., safe practices for cord prolapse, shoulder dystocia, and obstetric hemorrhage and safe cesarean section), units focused on clarifying and standardizing staff roles and responsibilities, and identifying and remediating latent system issues preventing a timely or appropriate response. Key findings for each strategy based on the triangulation of implementation monitoring data and interviews are described below:

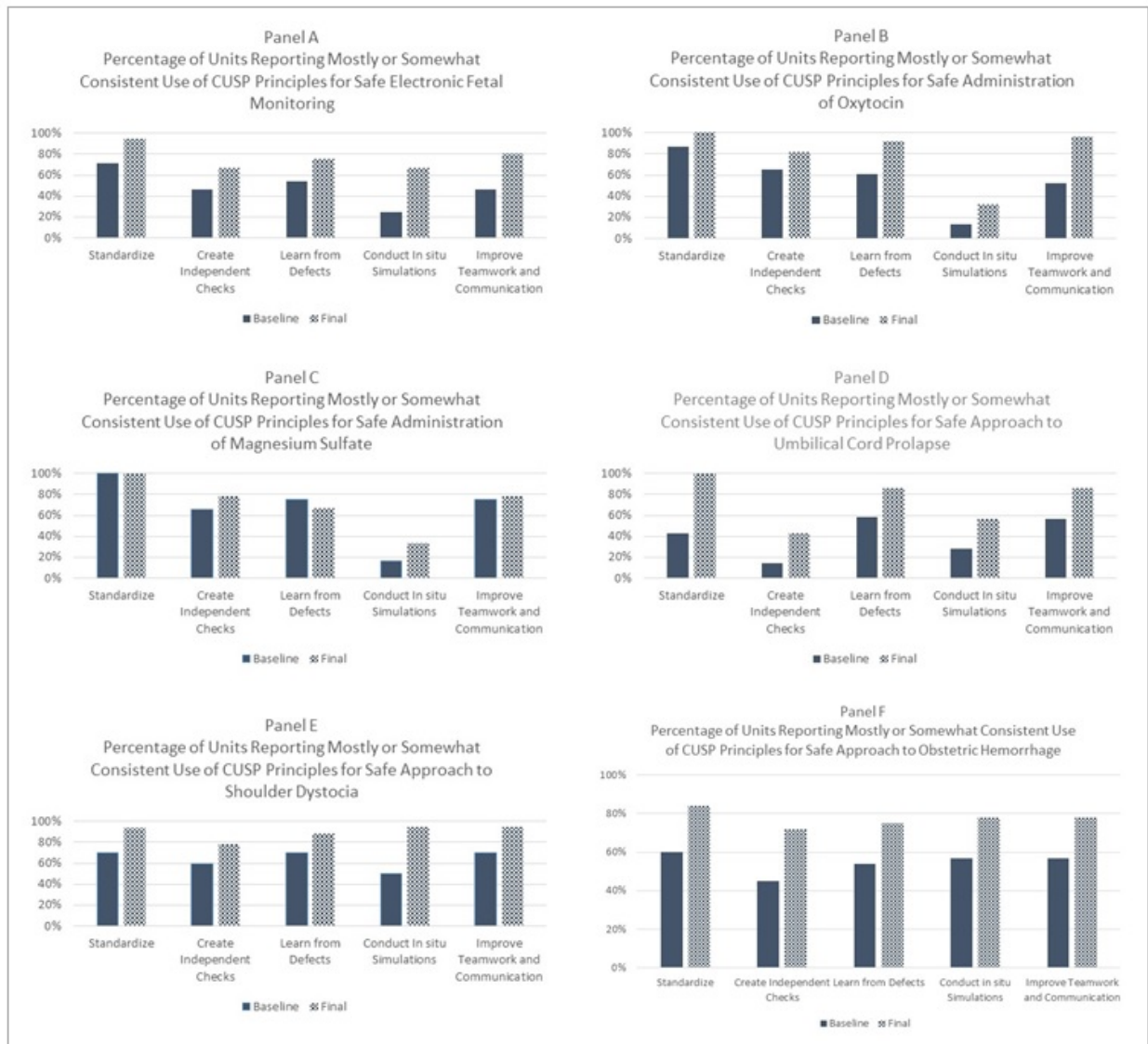
- Twenty-one L&D units elected to implement a strategy for **safe EFM**. The consistent use of all key Comprehensive Unit-based Safety Program (CUSP) principles increased from baseline to final followup for this strategy as indicated in **Panel A of Figure 6**. The mean percent of births with EFM that were documented using standard National Institute for Child Health and Development EFM nomenclature did not change appreciably (86 percent at baseline vs. 83 percent at final followup). Many units focused on making changes to their EFM training and/or certification protocols, updating existing policies with current EFM standards, and applying teamwork and communication techniques in communicating EFM tracings or strip reviews. These changes improved physician–nurse communication about EFM findings, built consensus on EFM strip interpretation, and resulted in policies that required more frequent EFM certification of L&D and antepartum unit staff.
- Twenty-five units elected to implement an obstetric **rapid response system**. The percent of L&D units with a standardized approach for 24-hours 7-days-a-week response to obstetric emergencies increased from 70 percent at baseline to 90 percent at final followup. Some units established new rapid response teams composed of obstetric providers, and others added obstetric nurses and physicians to the existing hospitalwide emergency response teams. Units also standardized the process for responding to obstetric emergencies in the emergency department and improved their communication with emergency department staff for coordinating this response. Teamwork and communication was vital to introducing or improving rapid response systems. L&D unit staff deployed techniques, such as closed-loop communication, situational awareness, and escalation of the chain of command. On some units, patients and families were also encouraged to activate a rapid response. Unit leaders observed that staff had increased situational awareness, familiarity with in-house resources, and communication between units during emergencies.
- Twenty-two L&D units elected to implement a strategy for the **safe administration of oxytocin**. The consistent use of all key CUSP principles increased from baseline to final

followup as indicated in **Panel B of Figure 6**, with largest increases in learning from defects and teamwork and communication. Units standardized oxytocin administration by drafting a joint oxytocin policy for nurses and physicians and creating or changing order sets to standardize doses. A few units incorporated their revised oxytocin management policy into their policies related to early elective inductions. Several units sought to ensure that the same intravenous (IV) pumps were used throughout the unit or purchased new “smart” pumps, and began using colored tubing or tubing labels to distinguish oxytocin lines from other IV lines. Several other units implemented checklists or printed reference cards for new oxytocin protocols. TeamSTEPPS techniques facilitated checklist implementation and effectively communicated appropriate oxytocin usage.

- Nine units elected to implement the **safe administration of magnesium sulfate**. At baseline, all of these units reported somewhat or mostly consistent use of the “standardize” key CUSP principle. Changes in other key CUSP principles for this strategy are indicated in **Panel C of Figure 6**. Several units instituted double checks when nurses initiated magnesium sulfate, changed dosage, or changed staff.
- Thirteen L&D units elected to implement a **safe surgery checklist during cesarean section**. The mean percent of all cesarean section births during which the safe surgery checklist was used increased from 26 percent at baseline to 78 percent at final followup. Some units elected to use the safe surgery checklist during cesarean sections in order to ensure the highest standard of care for the high-risk patients; others used it only for emergency cesarean sections. Units used simulations, formal education, and informal discussions during staff meetings to teach staff about using the safe surgery checklist. Some posted the checklist in the operating room, integrated the checklist into their emergency cesarean rapid response protocols, and used them to conduct debriefs after unscheduled cesarean sections.
- Seven units elected to implement safe practices for responding to deliveries complicated by **umbilical cord prolapse**. The consistent use of all key CUSP principles increased from baseline to final followup as indicated in **Panel D of Figure 6**. Key efforts for these improvements centered on teamwork and communication and in situ simulations. Debriefings were central to understanding and improving such processes, because response to this emergency involved staff from multiple units.
- Eighteen units elected to implement safe practices for responding to deliveries complicated by **shoulder dystocia**. The consistent use of all key CUSP principles increased from baseline to final followup as indicated in **Panel E of Figure 6**. All units that selected this strategy conducted in situ simulations to improve teamwork and communication skills. Some units also worked to standardize equipment, policies, documentation practices, and roles and responsibilities of care teams.
- Thirty-two units elected to implement safe practices for responding to patients with **obstetric hemorrhage**. The consistent use of all key CUSP principles increased from baseline to final followup as indicated in **Panel F of Figure 6**. The mean percent of births with an obstetric hemorrhage risk assessment documented on admission increased from 29 percent at baseline to 74 percent at followup. Further, the mean percent of patients with obstetric hemorrhage that had quantitative assessment of blood loss increased from 57 percent at baseline to 74 percent at followup. Lastly, the mean percent of patients with vaginal delivery who were administered a uterotonic agent during the third stage of labor increased from 91 percent at baseline to 95 percent at followup. Units achieved these changes by introducing routine hemorrhage risk assessment protocols and processes for quantifying

blood loss during hemorrhages, adding to or enhancing hemorrhage kits and carts, programming hemorrhage medication kits into automated dispensing cabinets, instituting or improving the frequency of debriefs after hemorrhages, practicing new protocols and equipment with in situ simulations, and improving role clarity during hemorrhage episodes. Training and education constituted a large component of the implementation approach, particularly for quantitative assessment of blood loss.

Figure 6. Implementation of key CUSP principles, strategy-specific teamwork and communication, and in situ simulations for selected perinatal safety strategies, AHRQ SPPC, 2015–2016



4.1.3 In Situ Simulations

As reported through implementation monitoring data, the proportion of units that had fully or mostly established in situ simulations with multidisciplinary participation increased from 29 percent at baseline to 58 percent after 10 months of followup. Similarly, the proportion of units that had fully or

mostly established regularly scheduled simulations and that monitored participation in simulation training increased from 31 percent to 55 percent.

As unit staff shared during interviews following 10 months of implementation, the focus on teamwork and communication during in situ simulations was a new approach to most participating L&D units. Most units had some experience conducting “code blue” simulations (laboratory or in situ) that focused on assessment and improvement of clinical competencies; these simulations did not address teamwork and communication. Prior simulations were also not conducted regularly, and those that were held involved only a limited group of staff, primarily from one discipline (e.g., nurses or resident physicians). Units that had minimal experience with in situ simulations had to dedicate a significant amount of time to planning, scheduling, and facilitating the simulations. Typically, local facilitators and a small group of nurses and/or a hospital-based learning specialist led the planning and implementation of the in situ simulations. Units found the integration of teamwork and communication skills in simulations valuable, especially when compared with their previous experiences of simulations conducted in a laboratory setting. L&D unit staff reported that these trainings facilitated staff engagement, fostered culture change within the unit, and allowed staff to identify safety hazards in their actual environment.

4.1.4 Overall Implementation Effectiveness

Using the composite measure of implementation effectiveness described in **Section 3.2.3**, 55 percent (26 units) effectively implemented all three program pillars.

Results of the qualitative comparative analysis (QCA) revealed three different combinations of contextual features among L&D units with effective implementation. These combinations were as follows:

1. adequate resources, appropriate staff engagement, the absence of strong leadership support, and the absence of dedicated time for program implementation
2. strong leadership support, appropriate staff engagement, the absence of adequate resources, and the absence of dedicated time for program implementation
3. strong leadership support, adequate resources, and the absence of appropriate staff engagement

Over 90 percent of L&D units with any one of these combinations of features had effective implementation. However, these combinations had poor coverage; only 39 percent of L&D units with effective implementation exhibited one or more of these combinations of features. This means that other features not included in this analysis may be contributing to whether an L&D unit was able to effectively implement the Agency for Healthcare Research and Quality (AHRQ) Safety Program for Perinatal Care (SPPC). These findings also suggest that multiple pathways to achieving implementation effectiveness exist and that absences of one or more features may be offset by the presence of other features.

Two multivariate regression analyses were conducted to evaluate the relationship among multiple L&D unit features and effective implementation. In the first analysis, implementation effectiveness was evaluated using a multivariate logistic regression model that included the four contextual features described in the preceding QCA section (leadership support, adequate resources, dedicated time, and staff engagement). The model with these four features was significant overall ($p=0.04$) and explained

16 percent of the variation in the implementation effectiveness among L&D units. Nonsignificant trends for the independent association of adequate leadership support ($p=0.07$) and adequate resources ($p=0.07$) were identified.

In the second analysis, implementation effectiveness was evaluated using a multivariate logistic regression model that included categorical variables representing annual number of births (three categories), level of neonatal care (three categories), the baseline patient safety culture (patient safety grade <75 vs. ≥ 75), and participation of the L&D unit in a perinatal safety or quality initiative at baseline. The model with these four features was significant overall ($p=0.008$) and explained 29 percent of the variation in the implementation effectiveness outcome among L&D units. Nonsignificant trends for the independent association of higher number of births ($p=0.06$) and higher baseline patient safety culture scores ($p=0.10$) were identified.

Common implementation challenges reported during interviews conducted 10 months following implementation that were relevant to all three pillars included the lack of staff time for activities to improve quality and safety in addition to existing clinical activities, difficulties in engaging physicians, and demands of competing local, State, and national initiatives.

4.2 Impact Evaluation Findings

The impact evaluation findings were assessed using the CUSP Team Checkup Tool, exit interviews describing L&D unit staff perceptions of patient safety culture changes, and claims data submitted at baseline and at 10 months postimplementation to identify changes in the incidence of adverse events. The overall impact of the AHRQ SPPC cannot be attributed to any single program pillar; rather, the impact on the unit patient safety culture and incidence of adverse events is a result of all three pillars working together and mutually reinforcing each other.

4.2.1 *Impact on Unit Patient Safety Culture*

Before program implementation, the mean percent of units with staff ratings of the overall patient safety grade based on the AHRQ Hospital Survey on Patient Safety Culture as excellent or very good was 74 percent, slightly below the national obstetric unit benchmark of 76 percent.⁶ Composite scores on other dimensions of this survey ranged from a low of 42 percent for “Nonpunitive Response to Errors” to 73 percent for “Supervisor/Manager Expectations and Actions Promoting Patient Safety.” By design, a followup unit patient safety culture survey at 10 months postimplementation was optional. Throughout the implementation phase, units used the CUSP Team Checkup Tool to gauge the changes in the patient safety culture related to the use of a CUSP approach; improvements in all items in the “Knowledge/Skills” and “Attitudes/Beliefs” domains of this tool were observed. Findings in the “Resources” domain of this tool were mixed or unchanged over the course of implementation.

As reported during interviews following 10-month implementation, many L&D units observed numerous positive impacts on their unit culture, which led to safer clinical care processes. Staff of many units observed that their unit culture began shifting from a traditional hierarchical model to a team-based

approach that equally valued and engaged all unit clinicians. These changes were most notably observed through nurse empowerment and ability to challenge authority, effective engagement of physicians, and increased peer-to-peer support. Although the program was limited to L&D, for some units this culture shift rippled to other hospital units and, in some cases, other hospitals that were part of the same health care systems. Numerous L&D units noted improved communication with other hospital units, improved staff satisfaction, lowered turnover, reduced work-arounds, and improved awareness of safety issues.

4.2.2 *Impact on Adverse Events*

Absolute changes in adverse events from baseline to 10 months were measured using hospital discharge abstracts. These findings are provided in **Table 3**. For all adverse events and composite measures evaluated, a favorable impact is represented by a decrease in the incidence. Because baseline incidence for some of these events was low, small absolute changes resulted in large relative changes for some events. The modified Adverse Outcome Index and Modified Weighted Adverse Outcome Score, both composite measures of maternal and neonatal adverse events, decreased from baseline to followup, though these decreases were not statistically significant. Two of the three maternal adverse event measures (Patient Safety Indicator [PSI] 19 and Inpatient Quality Indicator [IQI] 33) had statistically significant decreases, while the third (PSI 18) also had a nonsignificant decrease. Two of the four neonatal adverse event measures (PSI 17 and the National Quality Forum-endorsed moderate unexpected newborn complications) had statistically significant increases in incidence. The increase in the total unexpected newborn complication measure was not statistically significant and was likely offset by a nonsignificant decrease in the severe unexpected newborn complication incidence.

Table 3. Summary of impact on maternal and neonatal adverse events for L&D units submitting adverse event data at baseline and followup (N=43) for the AHRQ SPPC, 2015–2016

Clinical Measure	Mean Percent (95% CI) ^a			
	Baseline ^b (95% CI)	Followup ^c (95% CI)	Absolute Change ^d (95% CI)	Relative Change
Modified adverse outcome index	0.050 (0.045 to 0.056)	0.047 (0.041 to 0.052)	-0.004 (-0.010 to 0.001) (P = 0.14)	-7.6%
Modified weighted adverse outcome score	1.36 (1.10 to 1.61)	1.33 (1.10 to 1.55)	-0.03 (-0.24 to 0.18) (P = 0.91)	-2.2%
Unexpected newborn complications (total)	3.99 (3.28 to 4.70)	4.25 (3.54 to 4.96)	0.26 (-0.27 to 0.80) (P = 0.19)	6.5%
Unexpected newborn complications (moderate)	2.06 (1.56 to 2.55)	2.54 (1.95 to 3.14)	0.49 (0.18 to 0.80) (P = 0.01)	23.8%
Unexpected newborn complications (severe)	1.94 (1.48 to 2.40)	1.70 (1.34 to 2.05)	-0.24 (-0.61 to 0.12) (P = 0.24)	-12.5%
Birth trauma (AHRQ PSI 17)^e	0.20 (0.12 to 0.29)	0.35 (0.21 to 0.50)	0.15 (0.06 to 0.24) (P < 0.001)	73.3%
Obstetric trauma vaginal deliveries with instrument (AHRQ PSI 18)^e	10.91 (8.70 to 13.11)	9.35 (7.07 to 11.63)	-1.56 (-4.18 to 1.05) (P = 0.11)	-14.3%
Obstetric trauma vaginal deliveries without instrument (AHRQ PSI 19)^e	1.87 (1.52 to 2.22)	1.42 (1.15 to 1.68)	-0.45 (-0.81 to -0.10) (P = 0.04)	-24.3%
Primary cesarean delivery rate: uncomplicated (AHRQ IQI 33)	17.61 (16.16 to 19.05)	16.49 (15.29 to 17.70)	-1.11 (-2.05 to -0.18) (P = 0.02)	-6.3%

^a Represents the mean percent of all eligible births or deliveries. Eligible births or deliveries vary by measure.

^b Mean percent for calendar year 2013 for most units, which corresponds to the period 12–18 months prior to AHRQ SPPC program implementation.

^c Mean percent for the 10-month period after the start of AHRQ SPPC program implementation.

^d Absolute change from baseline to followup, P value for statistical significance between baseline and followup measure calculated using Wilcoxon signed-rank test.

^e 1 unit that submitted data is excluded from all PSI adverse event measures because of discrepancies in PSI data caused by the transition from ICD-9 to ICD-10 diagnosis codes in the data source.

Abbreviations: AHRQ = Agency for Healthcare Research and Quality, L&D = labor and delivery, PSI = Patient Safety Indicator, IQI = Inpatient Quality Indicator

Some statistically significant differences in adverse events were observed when units were stratified by coordinating entity, hospital or L&D unit characteristic, perinatal safety strategy selected, baseline unit patient safety culture as measured by the AHRQ Hospital Survey on Patient Safety Culture, the unit's participation in other perinatal quality or safety initiatives at the time of enrollment, and by the unit's effectiveness of the program implementation. These findings include:

- More favorable changes in the total and the severe unexpected newborn complication measures for units with Level II or Level III neonatal intensive care units (NICUs) compared with units with a Level I NICU. Thus, the driver of overall increases in unexpected newborn

complications incidence across all sites appears to be the result of disproportionate increases in this adverse event on units with Level 1 NICUs.

- More favorable changes in the severe unexpected newborn complications measure for units with resident physicians providing care compared with units without residents providing care; further, larger numbers of residents on a unit were associated with more favorable changes for this adverse event.
- Less favorable changes in the obstetric trauma without instruments measure (PSI 19) for hospitals that were publicly owned compared with those that were not publicly owned.
- Less favorable changes in the birth trauma measure (PSI 17) for hospitals designated as rural health clinics compared with hospitals not designated as rural health clinics.
- More favorable changes in the obstetric trauma with instruments measure (PSI 18) for units with a higher number of L&D unit beds.
- More favorable changes in the severe unexpected newborn complication measure for units with higher proportion of nonwhite patients.
- More favorable changes in the total and severe unexpected newborn complication measure for units with a higher proportion of vaginal births after cesarean at baseline; similarly, less favorable changes in the severe unexpected newborn complication measure for units with a higher proportion of deliveries by cesarean section at baseline.
- Up to eight strategies were possible for selection, and selecting fewer strategies was associated with more favorable changes in the Primary Cesarean Uncomplicated Rate (IQI 33, $p=0.02$). Larger decreases were also seen in this adverse event measure by units that did NOT select the cord prolapse or shoulder dystocia strategies.
- No difference in any adverse events based on the unit's baseline patient safety culture.
- No difference in any adverse events based on whether the unit was participating in an existing perinatal quality or safety initiative at baseline.
- No difference in any adverse events based on how effectively the unit implemented the program.

5. Discussion

5.1 Summary of Findings

5.1.1 Overall Program Implementation

Nearly all participating units successfully progressed toward the implementation of one or more program pillars, and 55 percent effectively implemented all three program pillars.

Findings from the implementation evaluation demonstrated that Comprehensive Unit-based Safety Program (CUSP) can be successfully applied in labor and delivery (L&D) units and beyond the focus area of healthcare-associated infections where it has been mostly applied to date. The Agency for Healthcare Research and Quality (AHRQ) Safety Program for Perinatal Care (SPPC) is comprehensive in design because it is based on a foundation of teamwork and communication that is reinforced with systematically applied perinatal safety strategies that are further strengthened by the use of in situ simulations in the context of specific clinical situations. Further, safety science principles—such as standardizing processes, creating independent checks, learning from defects, and engaging patients and families—can be applied to routine obstetrical care processes and procedures for responding to urgent clinical situations. This design approach could be replicated to employ CUSP principles in other care settings.

Committed leadership, availability of adequate resources, and engagement of staff from multiple disciplines—including physicians, nurses, anesthesiologists, blood bank staff, and emergency department staff—facilitated overall program implementation. Units that used a top-down approach and mandated policy changes and trainings were most successful with implementation; those that were driven using a grassroots approach were less successful, as they often lacked representation of multiple disciplines on the implementation team, visible support from the leadership, or time and resources for implementation.

Although a comprehensive program design allowed units to customize program tools and define their own best pathways to effective implementation, it increased the complexity of the program. In turn, this high complexity hindered implementation and contributed to a large data collection burden, both of which resulted in high unit attrition during the implementation phase. Other implementation challenges centered around lack of unit staff time for quality and safety improvement, staff turnover, difficulties in engaging physicians, and demands of competing local and national priorities.

5.1.2 Teamwork and Communication

The AHRQ SPPC implementation experience supports findings from the literature that TeamSTEPPS® can be effectively implemented during routine or urgent clinical situations and institutionalized and sustained through practice of those skills during in situ simulations.⁷⁻⁹ At the end of the program, most units were trained in multiple TeamSTEPPS communication techniques. Units used different approaches to improve teamwork and communication, with many using a traditional train-the-trainer approach and relying on the TeamSTEPPS master trainers. Others utilized a condensed 2-hour TeamSTEPPS Webinar specifically geared toward the L&D environment, which allowed clinical staff to complete the training at a convenient time and did not require scheduling an entire group for training at the same time. Although units were guided to implement TeamSTEPPS as the first and necessary pillar before the other two pillars, an approach that limited the implementation focus to a few selected TeamSTEPPS techniques and allowed for the gradual introduction of additional techniques over time facilitated more successful implementation than an approach that relied on having master trainers complete full staff training on all TeamSTEPPS techniques. Extending TeamSTEPPS training to other units that L&D staff work with, such as neonatal intensive care unit (NICU) or mother/baby, may lead to continual reinforcement of the techniques as part of the hospitals' culture, lessening the reliance on single individuals (e.g., master trainers and champions) for long-term sustainment.

5.1.3 Perinatal Safety Strategies

Because this program pillar was designed to target a diverse range of obstetrical care practices and conditions, an overarching framework based on select CUSP principles was used to provide a cohesive implementation approach across the different strategies. Use of most CUSP principles increased from baseline to followup for nearly all strategies included in this pillar. However, few units reported consistently being able to engage patients and families, one of the CUSP principles. A combination of teamwork and communication skills specific to the strategy and reinforcement of these skills through in situ simulations effectively guided units' implementation activities. This pillar also offered L&D units the flexibility to select strategies and prioritize their improvement efforts based on their specific needs, which some units parlayed into an ongoing vehicle for unit improvements.

Several crosscutting and strategy-specific themes from this pillar emerged. These include the following:

- Strategies included in the AHRQ SPPC are not distinct and can be mutually reinforcing. For example—
 - Strategies related to safe medication administration of oxytocin and magnesium sulfate both included a focus on interpretation, communication, and documentation of electronic fetal monitoring (EFM) findings.
 - The rapid response system strategy provided an overall approach and role clarity for responding to urgent obstetric situations, while the condition-specific strategies (e.g., safe practices related to cord prolapse, shoulder dystocia, and obstetric hemorrhage) provided a tailored approach for specific conditions.

- The safe medication administration strategy was particularly amenable to the application of the CUSP principles of standardizing and creating independent checks.
- The strategy related to safe cesarean section, which involved the implementation of a safe surgery checklist, offered examples of the diverse ways in which a unit could customize its implementation approach.
- The safe practices for obstetric hemorrhage strategy was the most commonly selected strategy; many units selecting this strategy already had some momentum for implementation because of involvement in other perinatal quality or safety initiatives focused on this topic. Further, many external resources exist for this topic and could be leveraged to augment the tools available through the AHRQ SPPC.

5.1.4 *In Situ Simulations*

Findings from the implementation evaluation confirm existing literature indicating that in situ simulations offer a valuable platform for introduction, practice, maintenance, and sustainability of skills that advance teamwork and communication as well as clinical skills.^{10,11} Using in situ simulations to focus on the practice of teamwork and communication skills, in contrast to clinical skills alone, was a new approach for most participating L&D units. Although few units were able to fully implement all aspects of a comprehensive program, the AHRQ SPPC offered units a platform for moving from the occasional ad hoc simulation, typically held with only nursing or resident physician staff, to an institutionalized, multidisciplinary strategy for regular staff training. Simulation scenarios provided as part of the AHRQ SPPC toolkit were the most unique and valued set of program tools. With these scenarios as models, several participating L&D units successfully developed and used scenarios for additional clinical processes or obstetric emergencies within or outside of L&D units, thus demonstrating the value of simulation as an organizational improvement strategy.

5.1.5 *Limitations of the Implementation Evaluation*

Few reliable and validated measures of implementation exist; thus, most measures used for the implementation evaluation were developed specifically for this evaluation during the pilot phase and refined over time. Tracking, compiling, and reporting implementation data proved to be a challenge for many units. As a result, some simply did not report these data or reported conflicting data across reporting periods. In particular, staff turnover and the rotational nature of resident physicians and other trainees thwarted unit attempts to report accurate data on TeamSTEPPS training and participation in in situ simulations. Measures developed for the perinatal safety strategies pillar sought to capture whether a unit was using safety principles to drive changes in processes and policies as opposed to measures of adherence to specific evidence-based clinical interventions. This differs from other perinatal safety initiatives that measure adherence to specific clinical interventions, and from other AHRQ CUSP healthcare-associated infection projects, in which the focus is on improving one adverse outcome (infection) and consensus exists for the evidence-based clinical interventions for reducing this outcome. Because of this, measures were not easily operationalized into the units' electronic health record (EHR)

to facilitate automatic electronic data extraction for monitoring and reporting. Based on this evaluation, further development and refinement of implementation measures are needed.

5.1.6 Summary of Impact on Unit Patient Safety Culture

The impact of the program on unit patient safety culture was largely positive. A quantitative assessment of change in unit culture was not possible because of the limited duration of the implementation phase (the AHRQ Hospital Survey on Patient Safety Culture may not be sensitive to change over periods of less than 1 year; thus, this survey was optional at the end of program implementation and only two units chose to re-administer it at followup). However, changes on the CUSP Team Checkup Tool during the implementation phase and participant feedback during the interviews suggest a favorable impact on culture. Through this program, many L&D units began the shift from a traditional hierarchical model to one characterized by a team-based approach with equal value placed on all unit staff. Changes in unit culture were most notable through the empowerment of nurses and their ability and willingness to challenge authority, the effective engagement of physicians, and increased peer-to-peer support. These changes resulted in improved staff satisfaction, reduced turnover, reduced work-arounds, and an overall improvement in awareness and importance of patient safety. Although the program was limited to L&D, for some hospitals the culture shift rippled to other units and, in some cases, other hospitals that were part of the same health care system.

5.1.7 Summary of Impact on Adverse Events

Ten months after the start of program implementation, nonsignificant decreases in the modified Adverse Outcome Index (MAOI) and Modified Weighted Adverse Outcome Score (MWAOS), both composite measures that combine maternal and neonatal adverse events, were observed, suggesting a favorable impact. When considering the maternal and neonatal events separately, decreases were observed for maternal adverse events, while increases were observed for all but one neonatal event. Because most of the adverse events monitored are rare, absolute changes in these events can result in large relative changes; thus, more data over time may be needed to observe clinically meaningful decreases in the absolute incidence of these events. Few significant differences in adverse events were observed by hospital or L&D unit characteristics, perinatal strategy selected, baseline patient safety culture, unit participation in other initiatives at baseline, and degree of effectiveness of the unit's program implementation.

5.1.8 Limitations of the Impact Evaluation

The flexible design of the program allowed L&D units to customize the approach to improve teamwork and communication and establish in situ simulations, and to choose which specific perinatal safety strategies to implement. This approach presents a challenge to a robust impact evaluation because units implemented a variety of different strategies at different levels of intensity or penetration and over different periods of time. Although flexibility may have permitted more widespread adoption of

the program than would otherwise have been possible with a rigidly proscribed program, it makes the interpretation of impact findings challenging because the program was implemented differently in each unit.

This evaluation had no comparison group to rule out the influence of secular trends, concurrent programs or initiatives, or other interventions. For example, the National Perinatal Information Center has observed an upward trend in the unexpected newborn complications measure since 2010,^b which may partially explain the evaluation findings of an increase in that measure over the 10-month observation period.

The measures used for adverse events were limited by the fact that most adverse events are rare, may not be totally preventable, and may not be sensitive to teamwork training and communication or specific perinatal strategies over the time period for which these units were observed. For example, some maternal or neonatal intensive care unit admissions or deaths may reflect serious underlying medical conditions, poor access to prenatal care, or other factors. The period of observation following the start of implementation was only 10 months, and the first several months comprised startup activities, with little change to existing processes or policies. The literature suggests that culture change takes years.¹² Other studies of similar initiatives monitor outcomes over several years.^{10,13–17} Further, hospitals participating in the National Perinatal Information Center's quality analytic services that use the MAOI typically review trends over 4 years.^b Thus, the evaluation findings of favorable impacts on some, but not all, adverse events over a relatively brief period are encouraging.

Administrative data were used to determine adverse events because this minimized the data collection burden for unit staff and offered a systematic approach across units. However, the disadvantage to this approach is that administrative data sometimes contain coding errors, and when events are rare (e.g., maternal deaths), these errors can make a difference in the adverse event incidence. In order to ensure data accuracy, identifiable data were required to enable the national team to work with units to resolve any coding errors. In other contexts, occasional errors in the data are unlikely to make a difference in overall outcome, and de-identified data from EHR extracts or administrative claim sources may be sufficient for evaluation of impact.

Lastly, the major limitation to the impact evaluation was the nationwide transition from International Classification of Diseases (ICD)-version 9 to ICD-version 10 codes for administrative/claims data, which occurred on October 15, 2015—in the middle of the 10-month observation period for this evaluation. Thus, data used to calculate adverse events before program implementation used algorithms and data with ICD-9 codes, while the adverse events at followup used algorithms and data with ICD-10 codes. This transition represented a major revision to the diagnostic code categories and definitions, and though some crosswalking between ICD-9 and ICD-10 is possible and was done, they are not equivalent coding schemes and AHRQ has discouraged the direct comparison of Patient Safety Indicator incidence rates between the two versions.

^b Personal Communication from Donna Caldwell and Janet Muri at the National Perinatal Information Center, September 21, 2016.

5.2 Recommendations for Future Program Design, Implementation, and Evaluation

5.2.1 Program Design

Two of the three program pillars—teamwork and communication, and in situ simulation—clearly distinguish the AHRQ SPPC from other perinatal quality and safety initiatives and should continue to be considered as components of the program, given that failures in communication are the most common root cause of adverse maternal events.³ However, additional options for TeamSTEPPS training, including condensed or abbreviated versions, should be considered. Further, the more these two pillars can be further tailored to the specific clinical problems and safety issues faced within the L&D unit environment, the better the content of these pillars will resonate with and serve the intended audience of perinatal physicians and nurses. One suggested change for future consideration is for the rapid response systems strategy from the perinatal safety strategies pillar to move to the teamwork and communication pillar, because this strategy can apply to a variety of different obstetric emergencies and is largely focused on establishing a systematic teamwork and communication approach across the care team.

The perinatal safety strategies pillar comprised a diverse set of varied strategies that straddle perinatal quality, safety, and clinical management. Some were specific to selected obstetric conditions (e.g., cord prolapse, shoulder dystocia, obstetric hemorrhage), and others offered a strategy for safe processes related to common procedures (e.g., EFM, medication administration, cesarean section). A more focused set of perinatal safety strategies with a clear and direct connection to very specific adverse events is one direction for how this pillar could be redesigned. This approach would require clear consensus and an evidence base for the unit-based or individual clinical actions required to prevent the adverse event, a way to measure that those actions occurred or did not occur, and a way to measure the adverse event.

This pillar could also be redesigned by broadening the set of strategies that are included within it. For example, this pillar currently does not address quality or safety issues related to surgical site infection, maternal or newborn sepsis, venous thromboembolism, or pregnancy-induced hypertension. Broadening this pillar may also warrant consideration as to whether strategies relevant to L&D-related units, such as antepartum, postpartum, mother/baby unit, and NICU, should be incorporated. If this pillar is broadened, it will be important to organize, integrate, and communicate the strategies in a logical way to avoid redundancy, ensure appropriate target audiences, and prevent hospitals from encountering decision paralysis and improvement fatigue.

5.2.2 Program Implementation

Based on the implementation experience associated with this program, focusing further dissemination efforts to coordinating entities, health care systems, and individual hospitals that are not already participating in existing perinatal quality or safety collaboratives may offer the most benefit. Such

a strategy would mitigate the barriers experienced as a result of competing initiatives and quality improvement fatigue.

A larger program impact might be seen if future implementation efforts, particularly those related to teamwork and communication and in situ simulations, were extended to related units (e.g., NICU, antepartum, mother/baby units). Such an approach to implementation would further improve teamwork and communication before, during, and after L&D and may bring the maximum benefit to the units' patient safety culture and to patient outcomes.

Although the majority of participating units had effective implementation, a sizable proportion of units struggled to implement one or more program pillars. Implementation support provided by the national team was virtual, largely through group-based Webinar training sessions, individual telephone calls, and virtual contact via email. Program implementation might be more effective if face-to-face coaching with units was included as part of implementation support. Such coaching was integral to the success of AHRQ's Patient Safety and Medical Liability Reform Initiative grantees, a perinatal safety initiative that served as the model and evidence for the AHRQ SPPC.^{14,16,17} Face-to-face coaching may be of greatest benefit to units and hospitals that are new to in situ simulations or smaller hospitals with few internal resources for quality improvement activities.

Finally, the ability to learn from defects required identifiable data pertaining to adverse events; some units dropped out of the program over concerns related to releasing identifiable data outside of their organization to the national team. Thus, future dissemination efforts may want to limit participation to hospitals that either agree to share identifiable data externally at the time of enrollment or have resources to calculate their own adverse event rates; however, the integrity of a cross-unit evaluation may be diminished by the latter.

5.2.3 Program Evaluation

The major recommendation for future evaluation efforts associated with this program concerns the data collection and submission requirements. Thirty percent of the units that were initially recruited dropped out due to the data collection burden. Allowing units to focus on data collection that is most meaningful to them locally in order to drive improvement is one possible future approach, but this approach would hinder the ability to conduct a robust cross-unit evaluation. In a program with multiple potential strategies for implementation that can impact multiple adverse events, and where no additional resources are provided to units for data collection, expectations for comprehensive and rigorous data collection should be tempered.

Further development of implementation process measures is required for a more robust approach to monitor implementation at a unit level, and to understand how variations in implementation affect the patient safety culture and the incidence of adverse events. If the program design evolves to be more encompassing of additional perinatal safety strategies, it will be important to identify a limited set of key implementation measures for each strategy and limit data collection to those measures for a cross-unit evaluation. Consideration to designating some measures as *core* for a cross-unit evaluation, while

designating others as *local* to provide units with measures that they can tailor to monitor their own implementation efforts, may help to minimize the required data burden. Further alignment of AHRQ SPPC implementation measures with measures used by other perinatal quality and safety initiatives and that can be operationalized into the EHR for automatic electronic data extraction may also reduce the data burden.

Establishing a more robust evidence base for the teamwork training, in situ simulations, and rapid response systems in a future evaluation would strengthen the rationale for adopting these program components in future dissemination efforts, as they tend to require a significant investment of staff time and resources to be fully implemented. The use of stepped-wedge evaluation designs may be able to contribute to the evidence base for these individual components; however, these evaluation designs would have a direct impact on program implementation and would require different pillars and strategies to be implemented at specific times, with data collection coordinated very closely with the timing of implementation.

In order to observe an impact on maternal and neonatal adverse events, a larger unit sample size with an observation period of at least 2 to 3 years will likely be needed. This would allow sufficient time for units to fully implement all three program pillars and for culture change to happen. Monitoring adverse outcomes throughout implementation is important, but expectations about the magnitude and timing of improvements should be realistic for future dissemination efforts. If impact on adverse events is to be evaluated over a shorter term, then a data collection and evaluation approach focused on a narrower set of strategies for very specific adverse events may be the best option for demonstrating improvements over a shorter term.

Lastly, consensus on how to consider some care processes currently designated as adverse events is warranted. For example, maternal blood transfusion, a component of both the MAOI and MWAOS composite indicators of adverse events, is considered an adverse event that results from failing to rapidly detect, communicate, and provide appropriate care for an obstetric hemorrhage. However, a maternal blood transfusion may also be considered an appropriate and lifesaving intervention that may have otherwise prevented a maternal death, particularly within some clinical contexts; for example, patients with placenta previa in which bleeding may be part of the natural course of the condition.¹⁸ Similar concerns exist for other clinical care events captured by administrative data to define maternal or neonatal adverse events. Thus, further refinement of such measures in future evaluations is warranted.

6. References

1. Martin JA, Hamilton BE, Osterman MJK. Births in the United States, 2015. NCHS Data Brief No. 258. Hyattsville, MD: National Center for Health Statistics; Sep 2016. Available from: <https://www.cdc.gov/nchs/data/databriefs/db258.pdf>
2. Forster AJ, Fung I, Caughey S, et al. Adverse events detected by clinical surveillance on an obstetric service. *Obstet Gynecol*. 2006 Nov;108(5):1073-83. Epub: 2006/11/02. PMID: 17077227.
3. The Joint Commission. Sentinel event data: Root causes by event type 2004–2014. Oakbrook Terrace, IL: The Joint Commission; 2014. Available from: http://www.tsigconsulting.com/tolcam/wp-content/uploads/2015/04/TJC-Sentinel-Event-Root_Causes_by_Event_Type_2004-2014.pdf
4. American Academy of Pediatrics Committee on Fetus and Newborn. Levels of neonatal care. *Pediatrics*. 2012 Sep;130(3):587-97. Epub: 2012/08/27. PMID: 22926177.
5. Damschroder LJ, Aron DC, Keith RE, et al. Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implement Sci*. 2009 Aug;4(1):50. doi:10.1186/748-5908-4-50
6. Agency for Healthcare Research and Quality. Hospital Survey on Patient Safety Culture: 2014 User Comparative Database Report [Internet]. Rockville, MD: Agency for Healthcare Research and Quality; 2014 March. Available from: <https://www.ahrq.gov/professionals/quality-patient-safety/patientsafetyculture/hospital/2014/hosp14tableb-1.html>
7. Pasciolla S, Alba B. An interdisciplinary approach to creating a culture of safety. *J Obstet Gynecol Neonatal Nurs*. 2015 June;44(S1):S39. doi:10.1111/1552-6909.12721
8. Straub SD. Implementing best practice safety initiatives to diminish patient harm in a hospital-based family birth center. *Newborn and Infant Nursing Reviews*. 2010 Sep;10(3):151-6. doi:10.1053/j.nainr.2010.06.011
9. Lundeen S, Padilla M. Communication and collaboration: Creating a culture of teamwork through the use of TeamSTEPPS strategies. *J Obstet Gynecol Neonatal Nurs*. 2013 June;42(s1):S67-S68. doi:10.1111/1552-6909.12152
10. Riley W, Davis S, Miller K, et al. Didactic and simulation nontechnical skills team training to improve perinatal patient outcomes in a community hospital. *Jt Comm J Qual Patient Saf*. 2011 Aug;37(8):357-64. Epub: 2011/08/31. PMID: 21874971.
11. Miller KK, Riley W, Davis S, et al. In situ simulation: A method of experiential learning to promote safety and team behavior. *J Perinat Neonatal Nurs*. 2008 April-June;22(2):105-13. PMID: 18496069.

12. Bisognano M, Cherouny PH, Gullo S. Applying a science-based method to improve perinatal care: The Institute for Healthcare Improvement Perinatal Improvement Community. *Obstet Gynecol.* 2014 Oct;124(4):810-4. PMID: 25198257.
13. Mann S, Pratt S, Gluck P, et al. Assessing quality obstetrical care: Development of standardized measures. *Jt Comm J Qual Patient Saf.* 2006 Sep;32(9):497-505. Epub: 2007/11/09. PMID: 17987873.
14. Riley W, Meredith LW, Price R, et al. Decreasing malpractice claims by reducing preventable perinatal harm. *Health Serv Res.* 2016 Dec;51(suppl 3):2453-71. Epub: 2016/08/22. PMID: 27549442.
15. Hendrich A, McCoy CK, Gale J, et al. Ascension Health's demonstration of full disclosure protocol for unexpected events during labor and delivery shows promise. *Health Aff (Millwood).* 2014 Jan;33(1):39-45. PMID: 24395933.
16. Demonstration Grants Final Evaluation Report [Internet]. AHRQ Publication No. 16-0038-2-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2016. Content last reviewed May 2016. Available from: <http://www.ahrq.gov/professionals/quality-patient-safety/patient-safety-resources/resources/candor/demo-program-eval/demo-grants.html>
17. Santos P, Ritter GA, Hefele JL, et al. Decreasing intrapartum malpractice: Targeting the most injurious neonatal adverse events. *J Healthc Risk Manag.* 2015 Apr;34(4):20-7. PMID: 25891287.
18. American College of Obstetricians and Gynecologists, Society for Maternal–Fetal Medicine, Kilpatrick SK, et al. Severe maternal morbidity: Screening and review. *Am J Obstet Gynecol.* 2016 Sep;215(3):B17-B22. Epub: 2016/08/22. PMID: 27560600.

List of Acronyms/Abbreviations

AHA	American Hospital Association
AHRQ	Agency for Healthcare Research and Quality
CEO	chief executive officer
CFIR	Consolidated Framework for Implementation Research
CUSP	Comprehensive Unit-based Safety Program
doi	Digital Object Identifier
EFM	electronic fetal monitoring
EHR	electronic health record
ICD	International Classification of Diseases
ICU	intensive care unit
IQI	Inpatient Quality Indicator
IV	intravenous
L&D	Labor & Delivery
MAOI	modified Adverse Outcome Index
MWAOS	Modified Weighted Adverse Outcome Score
NCHS	National Center for Health Statistics
NICU	neonatal intensive care unit
PMID	PubMed ID
PSI	Patient Safety Indicator
QCA	qualitative comparative analysis/analyses
SBAR	Situation, Background, Assessment, Recommendation
SD	standard deviation
SPPC	Safety Program for Perinatal Care
VBAC	vaginal birth(s) after cesarean