A literature review of quantitative indicators to measure the quality of labor and delivery care

Vandana Tripathi *
Department of Population, Family, and Reproductive Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
EngenderHealth, New York, NY, USA

Background: Strengthening measurement of the quality of labor and delivery (L&D) care in low-resource countries requires an understanding of existing approaches. Objectives: To identify quantitative indicators of L&D care quality and assess gaps in indicators. Search strategy: PubMed, CINAHL Plus, and Embase databases were searched for research published in English between January 1, 1990, and October 31, 2013, using structured terms. Selection criteria: Studies describing indicators for L&D care quality assessment were included. Those whose abstracts contained inclusion criteria underwent full-text review. Data collection and analysis: Study characteristics, including indicator selection and data sources, were extracted via a standard spreadsheet. Main results: The structured search identified 1224 studies. After abstract and full-text review, 477 were included in the analysis. Most studies selected indicators by using literature review, clinical guidelines, or expert panels. Few indicators were empirically validated; most studies relied on medical record review to measure indicators. Conclusions: Many quantitative indicators have been used to measure L&D care quality, but few have been validated beyond expert opinion. There has been limited use of clinical observation in quality assessment of care processes. The findings suggest the need for validated, efficient consensus indicators of the quality of L&D care processes, particularly in low-resource countries.

© 2015 International Federation of Gynecology and Obstetrics. Published by Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Although the rate of maternal death has decreased globally, many low-resource countries will not achieve the Millennium Development Goal (MDG) 5 to reduce maternal mortality [1–3]. Similarly, despite reductions in the past two decades, 2 million intrapartum stillbirths and intrapartum-event-related early neonatal deaths occur each year [4].

Skilled birth attendance rate—a commonly used measure of progress toward reducing maternal mortality—is included in the list of MDG 5 indicators [1]. Although the rates of facility delivery and skilled birth attendance are increasing in many low-resource countries, service contacts are not informative about the quality of labor and delivery (L&D) services, including essential newborn care (ENC) [5,6]. The content and quality of care (QoC) are crucial in ensuring the provision of interventions that either reduce the incidence of intrapartum and postpartum complications or respond to these complications [6–8].

Thaddeus and Maine’s widely used “three-delays” framework of maternal mortality [9] explicitly links QoC to the first and third delay, and proposes that perceptions of quality could be more important than access and distance in the decision to seek care. Empirical research suggests that poor QoC could underlie persistently high maternal mortality despite increasing facility delivery [10,11]. Studies indicate that perceptions of poor QoC lead to both a low demand for facility-based L&D services and a bypassing of close-by facilities for more distant ones [12].

Despite the evident importance of L&D care quality in reducing mortality and morbidity, questions remain about how to define and measure this construct. Many definitions of QoC have been proposed, including the WHO description of quality as encompassing effective, efficient, accessible, acceptable, patient-centered, equitable, and safe services [13]. However, these comprehensive definitions need refinement to enable an assessment of L&D care. The Donabedian QoC framework is useful in conceptualizing L&D care assessment, identifying three components of quality—namely, structure, process, and outcomes [14].

The present review had three aims. The first was to identify, describe, and classify in accordance with the components of the Donabedian QoC framework, quantitative indicators that have been proposed or applied to assess the quality of facility-based L&D care, including during the intrapartum and immediate postpartum period, and ENC. The second was to describe how quality indicators were selected and the data collection approaches used to evaluate these indicators. Finally, the review sought to identify gaps in QoC indicators used currently that should be addressed through future research in low-resource countries.
2. Materials and methods

In a systematic review, the PubMed, Embase, and CINAHL Plus databases were searched to identify research on quantitative indicators of L&D care quality published in English between January 1, 1990, and October 31, 2013. The search terms combined the following words and phrases: “maternal,” “obstetric,” “newborn,” “L&D,” “QoC,” “performance,” “measure,” “indicator,” “process indicators,” “assessment,” and “standards.”

Journal articles identified through the searches were retained for full-text review when the citation or abstract suggested that the study contained a description of proposed or applied indicators of facility-based L&D care quality. Although the present review was motivated by concerns related to care in low-resource countries, articles were included from all countries under the presumption that many aspects of QoC are independent of context.

Articles were excluded after citation or abstract review if they referred solely to community practices or home delivery, prenatal care, care after the immediate postpartum period, and prevention of mother-to-child HIV transmission without reference to other aspects of L&D care. Dissertations, conference proceedings, and books were excluded.

Articles meeting the criteria received full-text review. Additional exclusion criteria were applied during full-text review to ensure a focus on the aims of the present review (Box 1).

Full-text review and abstraction of information from articles was conducted with a structured spreadsheet template in Microsoft Excel 12.0 (Microsoft Corporation, Redmond, WA, USA). The assessed article characteristics included study country or region, methods used to select and validate indicators, indicator data sources, inclusion of newborn care or maternal or newborn complications, and distribution across the components of the Donabedian framework [14].

The present review sought to represent the investigators’ descriptions of QoC indicators. For example, among the widely used UN process indicators for emergency obstetric and newborn obstetric care (EmONC), only facility-based case fatality rate is classified as a QoC indicator; other UN process indicators are described as measures of availability, access, and utilization [15]. In the present review, therefore, studies applying only the UN process indicators were categorized as using a single or sentinel QoC indicator, although it is possible to interpret service availability indicators as measures of structural quality [14].

The review was conducted in adherence with PRISMA guidelines [16].

3. Results

Fig. 1 summarizes the article search and selection process. Application of the structured search terms across three databases identified 1224 unique articles. After abstract and full-text review, 477 articles were included in the present analysis (Supplementary Material S1).

Table 1 summarizes key characteristics of articles included in the review, specifically geographic focus, indicator selection processes, data collection approach, inclusion of Donabedian QoC framework components, number and/or type of indicators (single/sentinel, composite, or multiple), inclusion of newborn care, and inclusion of complication care. Among the 477 articles included, studies were evenly split between high-income countries (HICs) and low- and middle-income countries (LMICs). Despite considerable overlap in the indicators proposed for HIC and LMIC settings, few studies described the indicators that they used as global or suitable for use across both HICs and LMICs. Overall, 170 (35.6%) articles used literature review in indicator selection and 147 (30.8%) referred to existing tools to identify indicators. All or some of the UN EmONC process indicators were used in 74 LMIC studies; these articles generally did not describe additional indicator selection processes. Several studies that applied UN process indicators without including the case fatality rate indicator were not included in the review.

### Box 1

Exclusion criteria applied during full-text review.

<table>
<thead>
<tr>
<th>Articles were not included in analysis if they exclusively described:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• A set of quality indicators without at least some illustrative examples of specific indicators</td>
</tr>
<tr>
<td>• Access to and availability of maternal and neonatal health services</td>
</tr>
<tr>
<td>• Adverse event reviews to identify substandard care without specification of QoC indicators or criteria (e.g. non-criteria-based clinical audit)</td>
</tr>
<tr>
<td>• Assessment of health systems capacity or service quality without a focus on intrapartum and immediate postpartum or neonatal care</td>
</tr>
<tr>
<td>• Care for induced abortion, ectopic pregnancy, or obstetric fistula</td>
</tr>
<tr>
<td>• Clinical guidelines or competence standards without explicit reference to their use as quality indicators</td>
</tr>
<tr>
<td>• Data sources and systems for QoC assessment without discussion or endorsement of specific quality indicators</td>
</tr>
<tr>
<td>• Extra-medical services (e.g. transportation and/or communication systems)</td>
</tr>
<tr>
<td>• Evidence for clinical procedures (e.g. active management of the third stage of labor or partogram) without endorsing specific indicators for quality in performing these procedures</td>
</tr>
<tr>
<td>• Indicators that were inappropriate, not feasible, or not meaningful for assessing QoC</td>
</tr>
<tr>
<td>• Indicators selected owing to their role in malpractice claims or healthcare costs</td>
</tr>
<tr>
<td>• Labor induction, pain management, or anesthesia without reference to overall labor and delivery care</td>
</tr>
<tr>
<td>• Maternal or newborn mortality levels without explicit identification of quality indicators</td>
</tr>
<tr>
<td>• Patterns of current clinical practices (e.g. cesarean rate, uterotonic administration, or partogram use) without explicit discussion of QoC</td>
</tr>
<tr>
<td>• Process or feasibility of quality assurance or improvement techniques (e.g. clinical audit) without discussion of specific indicators to measure quality</td>
</tr>
<tr>
<td>• Qualitative data collection or qualitative exploration of QoC without prespecified quality indicators</td>
</tr>
<tr>
<td>• Rates of obstetric complications, near misses, or severe morbidities without explicit identification of them as quality indicators</td>
</tr>
</tbody>
</table>

### Abbreviation: QoC, quality of care.

Many articles referred to clinical guidelines, professional association recommendations, and government policies in indicator selection. Several studies seemed to convert clinical guidelines into a set of indicators, particularly to assess management of complications. Numerous studies described the use of expert opinion ranging from informal staff committees to formal Delphi processes to select quality indicators. Some of these described providing experts with an explicit set of criteria for the selection process (e.g. availability in existing clinical data sources).

Indicator selection commonly involved two stages. The first was a review of published literature, clinical guidelines, and/or existing QoC tools. The second was a critical analysis of information gleaned through desktop review by a panel of experts, such as public health leaders, clinicians, or other health-system representatives. Few articles described including service users during the process of indicator selection (Table 1). Some selected quality indicators through empirical validation, such as examining the association of potential indicators with clinical outcomes or the correlation between performance of potential indicators and associated constructs. In general, studies conducting such validation applied literature review or expert opinion to identify the pool of potential quality indicators.
In some studies, the authors discussed the clinical importance of a problem or concern regarding current practices (e.g. rising cesarean rates), justifying related indicators as measures of quality. In some articles about health service strengthening projects, a priori program indicators were used to evaluate QoC. Only 24 (5.0%) articles did not discuss the context or criteria for indicator selection.

Most articles described data collection to apply QoC indicators and measure quality. The most common data collection method was review of existing medical data sources, such as patient records, hospital registries, and administrative and/or routine databases (Table 1). Interviews or surveys of facility managers and providers were frequently used, as were other on-site facility assessments, particularly to ascertain structural quality. Several studies elicited information from service users through surveys, focus groups, and exit interviews. Only 38 (8.0%) articles used direct service observation to collect data on quality indicators, although an additional 7 (1.5%) articles described observation of drills or simulations on anatomical models. Some articles recommended QoC indicators but did not apply them (74, 15.5%). The articles described indicators across the structure, process, and outcomes components defined by Donabedian [14]. QoC was most frequently assessed through indicators of maternal and neonatal outcomes and care processes. Indicators of structural quality (e.g. provider knowledge or facility readiness) were noted in only 119 (24.9%) articles.

Most studies proposed or applied multiple indicators of quality, although approximately one-fifth described a single/sentinel indicator. A few articles proposed or applied a composite measure, using multiple indicators to create one quality index or scale score.

Table 2 lists the 17 single/sentinel indicators used across the 477 studies in the review. Measures of maternal deaths (i.e. facility-based case fatality rate or maternal mortality rate/ratio) were the most frequent single indicator, probably because the case fatality rate is identified as the sole quality measure among the UN process indicators. The proportion or rate of cesarean delivery was also applied as a single/sentinel indicator of quality, generally in HICs, where higher rates were viewed as a marker of poor quality. Perinatal mortality and its components (stillbirth, intrapartum stillbirth, and early neonatal death) were also applied as single/sentinel indicators. Other single indicators of quality described in multiple articles included severe maternal morbidities or maternal near miss, severe perinatal tear rate (or, conversely, rate of intact lower genital tract), patient satisfaction, and interval either between admission and provision of care or between the decision about and delivery of an intervention. Only one article proposed a structural measure (staff availability) as a single/sentinel quality indicator (Table 2).

Table 3 identifies the 15 composite measures that were developed or applied in the studies. Six composite measures exclusively reflected outcomes such as adverse clinical events or client satisfaction. For example, the Adverse Outcomes Index assessed the proportion of
deliveries in which any of 10 maternal or newborn events occurred (e.g. third- or fourth-degree perineal tears or neonatal death) [17]. Six composite measures exclusively assessed care processes. For example, the Bologna Score consists of five actions to be taken in the intrapartum/immediate postpartum period (e.g. partogram use and skin-to-skin care contact between the mother and newborn) [18]. Two composite measures included both processes and outcomes; one composite measure exclusively assessed structural quality. The composite measures are fully described in Supplementary Material S2.

For quality assessment, 355 (74.4%) articles used multiple indicators, ranging up to several hundred in some cases and often including both care processes and clinical outcomes. Supplementary Material S3 describes these sets of indicators for articles from LMIC settings. In general, articles proposing multiple indicators for separate evaluation (as opposed to combination into a single composite measure) did not rank, weight, or otherwise prioritize indicators. Approximately half (252 [52.8%]) of the articles included at least one quality indicator related to the newborn (e.g. readiness to provide

### Table 2

Proposed sentinel or single indicators of intrapartum and immediate postpartum QoC.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Variations, subgroups, and adjustment variables</th>
<th>No. of articles</th>
<th>Structure</th>
<th>Process</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case fatality rate/maternal mortality rate</td>
<td>Direct obstetric case fatality rate With and without complications Cause-specific Maternal mortality ratio Dichotomized at mean Cause distribution Timing distribution Preventable cesarean rate Proportion of deliveries that are cesarean Risk-adjusted Risk-adjusted and dichotomized into lower or higher than expected Ratio of actual to predicted risk-adjusted cesarean rates Unjustified cesarean rate</td>
<td>41</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cesarean delivery rate&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Elective repeat cesarean rate Nulliparous, term, singleton, vertex cesarean rate, adjusted for clinical and non-clinical (i.e., demographic) variables Nulliparous singleton vertex cesarean rate, adjusted for maternal age Optional vaginal delivery rate Preventable cesarean rate Proportion of deliveries that are cesarean Risk-adjusted Risk-adjusted and dichotomized into lower or higher than expected Ratio of actual to predicted risk-adjusted cesarean rates Unjustified cesarean rate</td>
<td>15</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree/rate of damage to the lower genital tract</td>
<td>Intact lower genital tract 3rd- or 4th-degree perineal tear rate Adjusted for instrumental assistance, epidural analgesia/anesthesia, and total vaginal deliveries per annum at facility Risk-adjusted for maternal age, parity, race, instrumental assistance, episiotomy, birthweight, and shoulder dystocia</td>
<td>3</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ideal delivery rate (absence of any of 26 adverse maternal/fetal outcomes)</td>
<td>–</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor induction rates</td>
<td>Primigravida</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neonatal near miss rate</td>
<td>–</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obstetric infection rates</td>
<td>Risk adjusted</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient satisfaction</td>
<td>–</td>
<td>2</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perinatal infection rate</td>
<td>Adjusted for case mix Disaggregated into maternal and neonatal Risk-adjusted</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perinatal mortality rate</td>
<td>Intrapartum and day 1/very early neonatal mortality rate Neonatal mortality rate Intrapartum stillbirth rate Adjusted for population variables Adjusted for case mix Classified by Nordic-Baltic system Crude and standardized based on birthweight and severity of illness Disaggregated into stillbirth, early neonatal mortality, and perinatal mortality Disaggregated into stillbirth and early neonatal mortality, compared by clinical risk groups Disaggregated by Wigglesworth criteria Owing to birth asphyxia</td>
<td>16</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peripartum hysterectomy risk by Robson group</td>
<td>–</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternal complication rate</td>
<td>Risk-adjusted Disaggregated by vaginal and cesarean delivery (different complications for each)</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe maternal morbidity/major morbidity/near miss rate</td>
<td>Condition on presentation at referral facilities</td>
<td>5</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff availability</td>
<td>Decision-to-incision time for emergency cesarean delivery Presentation-to-intervention interval Admission/decision to surgical intervention time</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time interval to receipt of care</td>
<td>–</td>
<td>3</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Umbilical pH &gt; 7.05 at delivery among newborns with a 5-min Apgar score &gt; 6</td>
<td>–</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacuum cup placement during vacuum-assisted delivery</td>
<td>–</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: QoC, quality of care.

<sup>a</sup> Different authors classify cesarean delivery rates as process or outcomes, depending on their orientation to QoC assessment. On the basis of the conceptual definitions of QoC applied in the present review and relevant literature from low- and middle-income countries, cesarean delivery rates have been classified as process indicators in the present analysis.
terns continue to blur the lines between LMICs and HICs. Second, it
picture of quantitative QoC indicators used in peer-reviewed studies
relied on existing data sources, primarily medical records.

tion of quality indicators. To collect data on QoC indicators, most studies
opinion. Less than 5% of studies included empirical or statistical valida-
and existing tools. Many investigators also referred to clinician or expert
studies used approaches such as literature review, clinical guidelines,

dicators of care processes and outcomes, rather than the structure com-
verse outcomes were included in most of the 477 studies (Table 1).

Abbreviations: QoC, quality of care; HIC, high-income countries; LMIC, low- and middle-income countries.

4. Discussion

The present literature review identified 477 articles describing quantitative indicators to assess the quality of intrapartum and immediate postpartum care, including ENC. The findings illustrate the diversity of indicators used to measure L&D QoC and demonstrate a lack of consensus regarding indicators of quality. Most studies emphasized indicators of care processes and outcomes, rather than the structure component of the Donabedian QoC framework. In selecting indicators, most studies used approaches such as literature review, clinical guidelines, and existing tools. Many investigators also referred to clinician or expert opinion. Less than 5% of studies included empirical or statistical validation of quality indicators. To collect data on QoC indicators, most studies relied on existing data sources, primarily medical records.

The present review has strengths. First, it provides a comprehensive picture of quantitative QoC indicators used in peer-reviewed studies from HIC and LMIC settings. This global scope is useful because rapidly changing economies, health system capacity, and service utilization patterns continue to blur the lines between LMICs and HICs. Second, it examines indicators relevant to both maternal and neonatal care. Such integration is crucial because, in many low-resource settings, one provider is responsible for both mother and newborn.

The review has some limitations related to exclusion criteria. Various terms are used by investigators to describe topics related to healthcare quality (e.g. performance, highly reliable organizations, competence, effectiveness, and safety). Interpretation and restriction were applied during the review to maintain a focus on intrapartum and immediate postpartum care and retain studies that could be analyzed in depth. Articles that did not explicitly refer to QoC were not included even if they examined related issues. As a result, the review does not cover all peer-reviewed studies relevant to L&D quality assessment. In addition, references in the included articles were not used to identify further studies for potential review. However, examination of references from a sample of articles found that most were already included in the review, and that they primarily described indicators that had been identified through the included articles. Assessments of care provision that did not use prespecified quantitative indicators, including numerous reports from death reviews and audits, were also excluded. Many of these studies identified “preventable deaths” or “avoidable factors” in broad categories (e.g. provider failures) that might inform the development of quantitative indicators for prospective quality assessment in the future. Moreover, much of the description of quality indicators is found in publications from professional associations and other non-peer-reviewed “grey” literature.

Another limitation is that the reliance on authors’ terminology had the potential to distort findings regarding the scope of quality indicators. Many investigators assessed characteristics that might be considered structural quality measures (e.g. staffing) without classifying these as quality indicators [19]. Several studies described the challenge of poor...
records in quality assessment [20]; however, few identified record-keeping as a quality indicator. Potentially relevant studies on continuity of care and midwifery teams were excluded when they did not characterize their aims as assessing care quality [21]. Such absence of QoC terminology might reflect differences in emphasis within provider cadres. Similarly, many studies examining women’s experiences of care were not included in the review because they tended not to include explicit, prespecified QoC indicators [22]. Such exclusion might have resulted in under-representation of interpersonal or respectful care as a quality indicator.

Overall, the exclusions and specific scope of the present review could have affected the numeric results. Several of the QoC indicators identified have been applied in more studies than are described here. However, the exclusions do not seem to have significantly affected the content of identified quantitative indicators. For example, many authors have reported findings on active management of the third stage of labor or the intrapartum and early neonatal death rate, but do not describe these studies as QoC assessments. Although such studies were not included in the present review, both of these aspects of care are described as quality indicators within the studies that were included, and thus are reported in the present findings.

The present review found that few indicators of L&D care quality have been assessed beyond face validity and, sometimes, content validity. Indicators have often been selected on the basis of clinical guidelines. Reference to evidence-based guidelines provides a level of validity but it does not guarantee that these practices are the best markers of overall QoC or that they facilitate practical differentiation of good or poor care. Furthermore, indicators derived from guidelines might not cover all dimensions of QoC. L&D care quality is a complex construct that probably involves more than just provision of evidence-based interventions. Souza et al. [23] recently found that high coverage with essential interventions does not imply reduced mortality, proposing that other services and overall improvements in care quality are required.

Indicators have frequently been selected on the basis of opinions of expert groups. Although this provides some validity, the composition of expert groups is often similar to a convenience sample comprising, for example, existing staff or members of professional associations. In addition, the process by which expert groups have selected indicators is not always explicit or systematic. Some studies described a set of criteria for selecting indicators, including availability in routine data sources, measurability, universal applicability, feasibility, amenability to change, and brevity [17,24,25], whereas other studies did not indicate how experts prioritized indicators. The validity of Delphi processes rests on standard procedures, such as selecting members representing an appropriate range of expertise and with access to up-to-date scientific evidence [26]. Few (2.3%) studies included service user perspectives in the selection of quality indicators, although health services research suggests that patients could prioritize attributes of care that are different to those targeted by providers and managers [27].

Because maternal complications are unpredictable and adverse outcomes are relatively rare, assessment of the process component of the Donabedian framework might be the most informative for quality improvement in maternal and newborn health care. The present examination of indicators used to evaluate L&D care processes illustrates the complexity of this task and the need for improved tools. Many studies of care process quality in low-resource countries have been criterion-based audits of adverse events, applying the approach described by Graham et al. in 2000 [8]. In general, criterion-based audits have assessed management of maternal complications such as postpartum hemorrhage, obstructed labor, and pre-eclampsia/eclampsia, sometimes directly applying national clinical guidelines as quality indicators.

Some sets of indicators to assess overall L&D care processes have included up to hundreds of items, partly because many process quality assessment tools are essentially checklists derived from clinical guidelines. However, clinical guidelines are meant to be exhaustive descriptions of care processes rather than selective indicators of quality. Chen et al. [28] noted that practice guidelines are qualitative recommendations, whereas quality indicators must be quantifiable and simple to measure. Conflict between the aims of guidelines and those of indicators could contribute to challenges in care process quality assessment [29].

In the present review, most studies assessing care processes relied on retrospective reviews of patient records, registers, and databases. However, medical records in low-resource settings are often incomplete [30,31]. Other commonly used data collection approaches also have limitations. Patient feedback might be subject to a “courtesy bias,” and women might be unable to provide accurate reports of many interventions during L&D care [32,33]. Studies have investigated whether population-based surveys can provide meaningful information on quality and coverage of essential interventions; initial research suggests that collecting data at facilities will remain essential [33,34].

On the basis of the present findings, several recommendations can be made for improving quality assessment through evaluation of quantitative indicators for intrapartum/immediate postpartum care. These recommendations are intended to facilitate ongoing quality assessment at the program level (i.e. at high-volume health facilities). These recommendations are also primarily intended for the low-resource setting. Although many aspects of QoC might be similar in low- and high-resource settings, the far greater mortality/morbidity burden and far more limited resources for quality assurance in low-resource countries require context-specific prioritization of quality indicators.

First, the burden of observation-based assessment of care process quality should be reduced. Research suggests that record reviews are weaker than observation in quality assessment, because observation can capture actions that are performed but not noted in medical records [31]. However, less than 5% of the reviewed articles used direct clinical observation to assess intrapartum/immediate postpartum care quality, and only seven used observation of simulated care. One possible solution is the identification of a few core indicators representing care processes that address causes of maternal and newborn mortality and morbidity, and that have been empirically validated and documented to be reliable. Bailit [35] and Mainz [36] provide the following criteria for such indicators: they should be evidence-based and associated with meaningful maternal and neonatal outcomes, related to outcomes influenced by provider/health system actions (i.e. amenable to change), easy to measure and observe reliably and reproducibly across settings, effective at discriminating between good and poor care, acceptable to providers and patients as meaningful quality markers, and affordable for application on a large scale. The Appraisal of Indicators through Research and Evaluation criteria also provides guidance for indicator selection [37].

Even with concise tools, observation of care requires more resources than other approaches. Thus, there is a concurrent need to strengthen alternative measurement approaches for core indicators of care quality. Provider self-assessment is an alternative, although evidence on its validity is mixed and further research is required on self-assessment of L&D care processes [38]. Routine data sources are likely to remain the most commonly used approach to quality assessment. The ability of such data to provide meaningful QoC information must be improved, perhaps through integration of quality assessment into ongoing provider support and record-keeping systems [39]. To encourage sustainable, ongoing evaluation of core quality indicators, inclusion of such measures into health management information systems might also be a long-term solution.

Lastly, health programs and systems require technical support to enable L&D care quality assessment. The present review suggests that the two most commonly used tools for assessing maternal and newborn care in LMICs are the UN EmONC process indicators and criterion-based audits of maternal deaths and near misses [15,40]. Both approaches have been disseminated with user guidance that outlines indicators, training needs, data collection tools, and analysis processes. Such guidance will be required for core quality indicators that may be identified and validated in the future.
In conclusion, the present review demonstrates the huge volume of indicators in use for the assessment of L&D care quality. The findings also indicate gaps and challenges in available quality assessment indicators. Many existing sets of quality indicators, particularly for care processes, are very long and difficult to apply. Indicators of care processes are rarely measured through direct clinical observation—the gold standard in clinical quality assessment. Few indicators have been validated beyond expert opinion or evaluated against systematic criteria such as those proposed by Bailit [35] and Mainz [36], making it difficult to evaluate how useful they are for quality assessment. The large number of nonvalidated indicators could contribute to the lack of consensus about which to prioritize in program settings. Overall, the findings suggest a need to improve options and guidance for ongoing quality assessment at the program level—the maternal health community must identify and prioritize validated, streamlined, quantitative indicators to facilitate observation-based assessment of the quality of intrapartum and immediate postpartum and newborn care, particularly in low-resource countries.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ijgo.2015.07.014.

Acknowledgments

This work was supported by the United States Agency for International Development (USAID award #GHS-A-00-08-00002-000) through the Maternal and Child Health Integrated Project.

Conflict of interest

The author has no conflicts of interest.

References


[4] Mathai M. To ensure maternal mortality is reduced, quality of care needs to be monitored and improved alongside increasing skilled delivery coverage rates. BJOG 2011;118(Suppl. 2):12–4.


[8] Mathai M. To ensure maternal mortality is reduced, quality of care needs to be monitored and improved alongside increasing skilled delivery coverage rates. BJOG 2011;118(Suppl. 2):12–4.


