Development of an instrument to evaluate intrapartum care quality in Senegal: evaluation quality care

ADAMA FAYE1,2, ALEXANDRE DUMONT3, PAPA NDIAYE1 AND PIERRE FOURNIER2,4

1Institut de Santé et Développement, UCAD, Dakar, Senegal, 2University of Montreal, Montreal, Canada, 3Institut de Recherche Pour le Développement, UMR 216, Université Paris Descartes, Sorbonne Paris Cité, Paris, France, and 4University of Montreal Hospital Research Centre (CRCHUM), 3875 Saint-Urbain St, 2nd Floor, Montreal, Quebec H2W 1V1, Canada

Address reprint requests to: adama.faye@umontreal.ca, adamfaye94@gmail.com

Accepted for publication 13 January 2014

Abstract

Objective. To evaluate the reliability of direct observation for measuring intrapartum care and compare this method with clinical audits using objective criteria based on patients’ medical charts.

Design. Cross-sectional study, data collected by two independent evaluators.

Setting. Hospital in Dakar, Senegal.

Participants. Thirty consecutive intrapartum care episodes provided by midwives and the corresponding medical charts.

Outcome Measure. The presence or absence of each of twelve criteria selected on the basis of national and international norms for monitoring of labour and delivery (six criteria) and the immediate postpartum period (six criteria).

Results. For direct observation, the labour and delivery mean quality scores ranged from 5.34 to 5.77. In contrast, for the chart-based method, the scores ranged from 0.32 to 0.45. For postpartum care evaluated only with direct observation, the scores were also high (5.21–5.65). For direct observation, inter-evaluator agreement was high: kappa coefficients varied from 0.78 to 0.93 depending on the criterion (total score ICC = 0.74). For the chart-based method, inter-evaluator agreement was also high: 0.66 to 1 (total score ICC = 0.72). Comparison of the two methods showed strong differences by items and subscores.

Conclusion. Using direct observation, the quality of obstetric care was high for both the monitoring of labour and delivery and postpartum care. Both measurement instruments showed high reliability. The chart-based method underestimated the quality of care because of poor medical record documentation. Medical-record-based measurement may not be appropriate for the evaluation of the quality of obstetric care in Senegal and other low-income settings.

Keywords: quality, obstetric care, reliability, Senegal

Introduction

Maternal mortality is a major public health problem in sub-Saharan Africa, where it remains very high in comparison with Western countries [1]. The main causes of maternal deaths are complications of abortion, obstetric haemorrhage (pre- and postpartum), uterine rupture, eclampsia and prolonged or obstructed labour [2]. Most of these deaths are avoidable but difficult to predict before the onset of labour and delivery. High-quality intrapartum care may detect these complications early and enable timely treatment.

To reduce maternal mortality, two strategies are currently recommended: assisted delivery and a package of essential obstetric services provided by trained staff [3]. This package of services, referred to as emergency obstetric and neonatal care, has six components: parenteral administration of antibiotics, of uterotonic drugs and of anticonvulsants, manual extraction of the placenta, forceps- or vacuum-assisted delivery and manual evacuation of residual uterine products. In developed countries, the low maternal mortality rate is largely due to the timely identification and appropriate treatment of obstetric complications [4]. Unfortunately, in Africa, the quality of services provided to patients is very heterogeneous, resulting in high rates of lethality in health care [5].

Measuring the quality of obstetric services is an essential step in any strategy for improving health care interventions to reduce maternal mortality. Several instruments have been developed to measure quality of care in different environments, including obstetric services [6]. The main instruments are as follows: evaluation of providers’ knowledge and attitudes in
specific cases (vignettes), evaluation of care based on medical charts (clinical audits with objective criteria) and direct observation of service providers during an episode of care. Vignettes are a method of choice, but they more often measure what the provider is supposed to do, rather than what is actually done [7]. Clinical audits of medical records are increasingly being used as a method to evaluate the quality of obstetric care in developing countries. However, the validity and reliability of this method are uncertain, given the heterogeneity of the information available in medical charts [8]. Indeed, medical charts are often poorly completed and rarely archived [9]. Direct observation of care constitutes a gold-standard reference method [10], but this method has rarely been evaluated in obstetric services in Africa. In addition, direct observation may influence both provider and patient behaviours, artificially inflating quality scores [11].

The objective of this study was to evaluate the reliability of direct observation for measuring intrapartum care in Senegal and to compare this method with clinical audits using objective criteria based on patients’ medical charts.

Method

Study population

This study, conducted uniquely in Senegal, was part of the QUARITÉ trial project implemented in both Mali and Senegal [12]. The objective of that randomized trial was to evaluate the effectiveness of a set of strategies to reduce maternal mortality in referral hospitals in these two countries. This study was carried out at Abass Ndao Hospital in Dakar, Senegal. This hospital, which has a maternity unit and surgical facilities, is located in the city centre. It serves an urban population drawn primarily from the working-class neighbourhoods of Dakar. The maternity service includes 21 midwives, 4 gynaecologist-obstetricians and 2 interns (medical trainees) in gynaecology-obstetrics.

Patients who delivered in this hospital’s maternity unit between 21 and 29 March 2011 and who were monitored during labour and delivery by a midwife were included in the study. Patients who delivered by caesarean or whose deliveries were forceps- or vacuum-assisted were excluded, to ensure observation of only one category of personnel, the midwives. The six midwives on duty during the study period all agreed to participate in the survey. Seven episodes of care were observed per midwife. To limit observation bias (Hawthorne effect), the two first episodes for each midwife were not considered in the analysis. The midwives and patients who agreed to take part in the study signed an informed consent form. The sample consisted of 30 patients. The participants were selected randomly among women who delivered in a given facility during an 8-day period. They are representative of the users of the facilities included in our study.

Measuring the quality of intrapartum care

To evaluate the quality of intrapartum care, we developed a standardized questionnaire for direct observation of care (‘GRIlle d’Observation des Soins’—GRIOS) based on the norms and procedures used in Senegal and the clinical practice guidelines of the International Federation of Gynecology and Obstetrics (FIGO). The GRIOS was validated by two independent experts, and by a third expert in cases of non-agreement between the first two. The quality evaluation was focused on the monitoring of labour and delivery in the delivery room and on the care provided in the immediate postpartum period (the 2 h following delivery). We selected 12 criteria for quality of care: 6 related to monitoring of the patient and foetus during the active phase of labour and 6 related to immediate postpartum care (Table 1). Each criterion was measured by direct observation (Method 1) as well as by detailed review of the medical chart 2 h post-delivery (Method 2).

Data collection

The data were collected by two independent evaluators using the GRIOS to capture information on the quality of the care as observed and as noted in the chart, as well as patients’ characteristics. The evaluators were two midwives who did not work at Abass Ndao Hospital. We chose to use midwives as evaluators to ensure an adequate level of knowledge to correctly evaluate the actions taken by the midwives working at the hospital but also to facilitate the evaluators’ integration into the delivery room. The two evaluators received 1 d of training for better standardization of quality-of-care measurement. The GRIOS was tested beforehand on a series of five episodes of care to verify the evaluators’ level of comprehension and their appropriate use of the measurement instrument.

Data collection was done in two stages for each episode of care. The first stage involved observing the midwife and patient throughout the labour, during delivery and in the 2 h post-delivery, and then recording information related to the care quality criteria in the GRIOS. While the midwife-clinicians were advised beforehand of the objectives of the study, they did not know which days or clinical acts would be covered by the quality evaluation. The second stage of the survey consisted of a detailed analysis of the medical chart 2 h post-delivery to collect information on the quality criteria (results of clinical examinations or acts specifically noted) and on the patient’s characteristics, services received and the outcome of the delivery.

Analysis

For each of the 12 quality evaluation criteria, a score of 1 was assigned if the criterion was met, and a score of 0 if it was not. Then, two sub-total scores of up to six each were calculated for each episode of care, corresponding to the sum of all scores for the six criteria related to the monitoring of labour and delivery (maximum score = 6) and to the sum of all scores for the six criteria related to immediate postpartum care (maximum score = 6).

Inter-evaluator agreement was assessed for each of the two methods (direct observation vs. chart-based evaluation) by calculating the intraclass correlation coefficient (ICC) [13] for the
total score and the adjusted Cohen’s kappa coefficient [14–16] for each quality criterion. According to Landis’ and Koch’s classification [17], agreement is considered satisfactory when the kappa coefficient is >0.61. We also used the method developed by Bland and Altman to represent graphically the extent of agreement between the two evaluators [18]. Agreement was considered satisfactory when 95% of the variation in total scores between the two evaluators was within two standard deviations of the mean.

To compare the two methods’ results with respect to quality measurement, the adjusted Cohen’s kappa coefficient was calculated for each criterion and the ICC for the total score. The analyses were conducted using R software (R 2.15.1—The R Foundation for Statistical Computing).

Results

The average age of the parturients was 25.8 years (SD = 5). These were first deliveries in 20% of the cases, and 45% of the women had received no schooling. In 40% of the cases, the women had undergone more than three prenatal consultations. All had come to the hospital on their own (not referred by another health facility). The average age of the midwives was 30.7 years (SD = 8).

Comparison of measurement methods

Agreement between scores measured by direct observation and those measured by chart-based evaluation was low for both evaluators (adjusted alpha coefficient of <0.1). Using direct observation during labour and delivery, the monitoring maternal and foetal well-being were carried out in >85% of the cases, regardless of the evaluator (Table 2). In contrast, the results of clinical examinations (blood pressure, heart rate, uterine contractions and vaginal examination) were not systematically noted in the obstetric chart: in the charts reviewed, between 3 and 16.7% of the results of exams had been recorded, depending on the selected criteria and the evaluators (Table 3).

The reliability of each measurement method was high. Inter-evaluator agreement for monitoring of labour and delivery, measured using the adjusted kappa coefficient, ranged from 0.78 (routine vaginal examination) to 0.93 (routine blood pressure readings) for direct observation, and from 0.66 (maternal heart rate noted in the chart) to 1 (foetal heart rate, rhythm and intensity of uterine contractions) for chart-based evaluation.

The mean quality scores by evaluator were 11.42 points (1.18) for evaluator 1 and 10.55 points (2.08) for evaluator 2 (maximum expected number of points = 12) for all 12 criteria.

### Table 1  Evaluation criteria for quality of intrapartum care

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Measurement method in direct observation</th>
<th>Measurement method based on medical charts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitoring of labour and delivery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood pressure</td>
<td>Measured every 2 h</td>
<td>Noted in the chart at each exam</td>
</tr>
<tr>
<td>Temperature</td>
<td>Temperature measured every 2 h</td>
<td>Temperature noted in the chart at each exam</td>
</tr>
<tr>
<td>Maternal heart rate</td>
<td>Measured every 2 h</td>
<td>Noted in the chart at each exam</td>
</tr>
<tr>
<td>Foetal heart rate</td>
<td>Measured every 2 h, either by intermittent auscultation or by cardiotocography</td>
<td>Noted in the chart at each exam</td>
</tr>
<tr>
<td>Uterine contractions</td>
<td>Rhythm and intensity of contractions measured every 2 h</td>
<td>Rhythm and intensity of contractions noted in the chart at each exam</td>
</tr>
<tr>
<td>Vaginal examination</td>
<td>Carried out every 2 h</td>
<td>Cervical dilation, foetal descent, status of foetal membranes and colour of amniotic fluid are noted in the chart at each exam</td>
</tr>
<tr>
<td>Immediate postpartum period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxytocin bolus</td>
<td>Injection at foetal shoulder delivery or immediately after birth</td>
<td></td>
</tr>
<tr>
<td>Uterine massage</td>
<td>Fundal massage performed before placenta expulsion</td>
<td></td>
</tr>
<tr>
<td>Cord traction</td>
<td>Traction performed on the cord before placenta expulsion</td>
<td></td>
</tr>
<tr>
<td>Examination of the placenta</td>
<td>Placenta examined on both sides by the midwife</td>
<td></td>
</tr>
<tr>
<td>Examination of the mother</td>
<td>Examinations (blood pressure, heart rate, temperature and vaginal bleeding) carried out 2 h after birth</td>
<td></td>
</tr>
<tr>
<td>Examination of the newborn</td>
<td>Examination of the newborn (colour, respiration, tone and heart rate) carried out 2 h after birth</td>
<td></td>
</tr>
</tbody>
</table>

Faye et al.
evaluated by direct observation \( (P = 0.16) \). Inter-evaluator agreement for the evaluation of immediate postpartum care using direct observation was high for the clinical examination of the mother and newborn 2 h post-delivery, but lower for uterine massage (adjusted alpha coefficient = 0.36) and examination of the placenta (adjusted alpha coefficient = 0.43). None of the criteria selected to evaluate the quality of postpartum care had been noted in the medical charts. Comparisons of the evaluators’ mean scores separately for direct observation and chart review were not statistically significant. The ICC between the two evaluators varied between 0.72 and 0.74 for monitoring of labour and delivery, and less within two standard deviations (Fig. 2).

**Discussion**

We developed an instrument to measure the quality of obstetric care using direct observation during the monitoring of labour and delivery and in the immediate postpartum period, adapting this instrument to the low-income country context. The results of this study showed that the GRIOS instrument is reliable, as is chart-based evaluation. On the other hand, chart-based evaluation markedly underestimated the quality of care when compared with direct observation. The inter-evaluator agreement was satisfactory, which augurs well for using the GRIOS to measure quality improvement in obstetric services after interventions are implemented to change practices.

The quality score obtained from direct observation of deliveries is higher than that obtained from analysing medical charts \[19\]. Indeed, there is a significant gap between what the midwives are observed to do and what they record in the patients’ medical charts. This underestimation of the quality of care when measured from medical charts or hospital registers is largely explained by poor documentation in clinical records, a problem observed previously by others \[9, 10\]. Attempting to use these medical charts or registers, in their current condition, as data sources for measuring quality of care or for conducting epidemiological studies in developing countries is clearly problematic \[9, 10\]. This poor chart maintenance can also be understood in the context of a professional culture in which there is not yet much awareness of the importance of medical record-keeping \[9, 20\]. However, when comparing

---

**Table 2** Measurement of the quality of monitoring of labour and delivery according to the two measurement methods

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Method 1 direct observation Evaluator 1</th>
<th>Evaluator 2</th>
<th>Method 2 chart review Evaluator 1</th>
<th>Evaluator 2</th>
<th>Comparison of the two methods Evaluator 1 kappa</th>
<th>Evaluator 2 kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood pressure measured</td>
<td>30 (100)</td>
<td>28 (93.3)</td>
<td>2 (6.7)</td>
<td>3 (10)</td>
<td>0.008</td>
<td>0.005</td>
</tr>
<tr>
<td>Temperature measured</td>
<td>30 (100)</td>
<td>27 (90)</td>
<td>3 (10)</td>
<td>2 (6.7)</td>
<td>0.009</td>
<td>0.006</td>
</tr>
<tr>
<td>Mother’s maternal heart rate measured</td>
<td>29 (96.7)</td>
<td>27 (90)</td>
<td>1 (3.3)</td>
<td>2 (6.7)</td>
<td>0.007</td>
<td>0.008</td>
</tr>
<tr>
<td>Newborn’s maternal heart rate measured</td>
<td>29 (96.7)</td>
<td>27 (90)</td>
<td>3 (10)</td>
<td>3 (10)</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Uterine contraction measured</td>
<td>26 (86.7)</td>
<td>25 (86.7)</td>
<td>1 (3.3)</td>
<td>1 (3.3)</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>Routine vaginal exam done</td>
<td>28 (93.3)</td>
<td>27 (90)</td>
<td>4 (13.3)</td>
<td>5 (16.7)</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>Sub-total score</td>
<td>5.77 (0.67)</td>
<td>5.34 (1.56)</td>
<td>ICC = 0.74</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3** Measurement of the quality of immediate postpartum care according to direct observation

<table>
<thead>
<tr>
<th>Method 1 direct observation Evaluator</th>
<th>Inter-evaluator agreement kappa</th>
<th>Method 2 chart review Evaluator</th>
<th>Inter-evaluator agreement kappa</th>
<th>Comparison of the two methods kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxytocin bolus</td>
<td>25 (86.2)</td>
<td>20 (69.0)</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>Fundal massage</td>
<td>26 (96.7)</td>
<td>19 (63.3)</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>Cord</td>
<td>26 (86.7)</td>
<td>26 (86.7)</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Placenta</td>
<td>28 (93.3)</td>
<td>29 (96.7)</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>Examination of mother post-delivery</td>
<td>30 (100)</td>
<td>30 (100)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Examination of newborn post-delivery</td>
<td>30 (100)</td>
<td>30 (100)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Sub-total score</td>
<td>5.65 (0.75)</td>
<td>5.21 (1.01)</td>
<td>ICC = 0.40</td>
<td></td>
</tr>
</tbody>
</table>
several facilities, or for rapid evaluation, chart analysis has been shown to be necessary; hence, the need to improve the maintenance of medical charts, not only getting providers more deeply involved [21] but also by doing more frequent quality control. The use of new information technologies may also be conclusive [22, 23]. More generally, special emphasis should be given to data management, both in basic training and in continuing education for medical and paramedical personnel [20].

The direct observation of immediate postpartum care was not as reliable for measuring the occurrence of interventions to prevent obstetric haemorrhage. This measurement requires proximity in order to observe the midwife’s actions, and this may lead to biased observation between evaluators. The result may also signal a difference in interpretation between the two evaluators, which could be addressed by additional evaluator training to improve the instrument’s reliability.

Most instruments used to measure obstetric care have not undergone any validation [7]. The items used to develop the quality scores were selected based on norms and protocols in effect in Senegal [24], instruments developed in Africa [19, 25, 26] and guidelines of FIGO. These items were validated by two experts in the field of obstetrics in Africa. These factors gave the GRIOS strong content validity with a useful international perspective.

The study has limitations. The number of evaluators and patients were limited. The study dealt only with the observable aspects of the interaction between providers and patients. Other aspects of the technical dimension such as empathy and courtesy were not measured. Likewise, it did not measure the structural dimension, which would have considered the availability of human and material resources as well as resource organizations, which are important aspects of health care quality [7], particularly in developing countries, where there is often a lack of even minimum resources required to carry out basic interventions [27].

**Conclusion**

Measuring the quality of obstetric care is critically important as part of efforts to evaluate health care interventions targeting quality of services, especially in developing countries, where many interventions are developed to improve the health of pregnant women and newborns. The GRIOS is an instrument with good content validity and high reliability, providing a useful perspective for evaluating programs in developing countries.

**Authors’ contributions**

AF supervised all the data collection in Senegal, wrote the first draft of the manuscript with AD and undertook the statistical analysis. AF, AD, PF and NDP led the development of the
study from its inception. All authors read, improved and approved the final manuscript.

Acknowledgements

We thank all the medical and administrative staffs of all the participating centres for their valuable contributions to this study. Adama Faye is a fellow of the Leadership Program in Population and Health funded by the Bill and Melinda Gates Foundation.

References


189