

# UCLA

## UCLA Previously Published Works

### Title

Variability in Health Care Quality Measurement among Studies Using Service Provision Assessment Data from Low- and Middle-Income Countries: A Systematic Review.

### Permalink

<https://escholarship.org/uc/item/47f7r574>

### Journal

American Journal of Tropical Medicine and Hygiene, 103(3)

### ISSN

0002-9637

### Authors

Moucheraud, Corrina  
McBride, Kaitlyn

### Publication Date

2020-09-02

### DOI

10.4269/ajtmh.19-0644

Peer reviewed

**Variability in health care quality measurement among studies using Service Provision Assessment data from low- and middle-income countries: A systematic review**

Corrina Moucheraud ScD MPH<sup>1</sup>\*, Kaitlyn McBride BA<sup>1</sup>

<sup>1</sup> University of California Los Angeles, Fielding School of Public Health, Department of Health Policy and Management; Los Angeles CA USA

\* Corresponding author

**Keywords:** Quality of care; health care facilities; global health

**Word count:** 245 (abstract), 3396 (manuscript)

**Number of figures:** 2

**Number of tables:** 2

**Number of appendices:** 1

**Running head:** Quality of care measurement using SPA surveys

---

\* Corresponding author  
UCLA Fielding School of Public Health  
Department of Health Policy and Management  
650 Charles E. Young Dr. S.  
Los Angeles, CA 90095 USA  
T: +1-310-206-1185; F: +1-310-825-3317  
cmoucheraud@ucla.edu

## Abstract

Quality of care is essential for improving health outcomes – but heterogeneity in theoretical frameworks and metrics can limit studies' generalizability and comparability. This research aimed to compare definitions of care quality across research papers that incorporate data from Service Provision Assessment (SPA) surveys. Following PRISMA guidelines, we used a keyword search in PubMed. Each author reviewed abstracts, then full texts, for inclusion criteria: peer-reviewed publications of empirical analysis using SPA data. The search yielded 3250 unique abstracts and 34 publications were included in the final analysis. We extracted details on the SPA dataset(s) used, theoretical framework applied, and how care quality was operationalized. The 34 included papers used SPA data from 14 surveys in 9 countries (all in sub-Saharan Africa plus Haiti). One-third of these papers (n=13) included no theoretical or conceptual framework for care quality. Among those papers referencing a framework, the most common was the Donabedian model (n=7). Studies operationalized quality constructs in extremely different ways. Few papers included Outcomes as a quality construct, and the operationalization of Structure varied widely. A key asset of SPA surveys, owing to the standardized structure and use of harmonized data collection instruments, is the potential for cross-survey comparisons. However, this is limited by the lack of a common framework for measuring and reporting quality in the existing literature using SPA data. SPA surveys offer unique and valuable insights, and a common framework and approach would substantially strengthen the body of knowledge on quality of care in low-resource settings.

## 1 Introduction

2 Quality of care is necessary for improving health outcomes<sup>1</sup>. Low-quality health systems may be  
3 responsible for up to 8.6 million deaths annually in low- and middle-income countries: i.e., these deaths  
4 could have been averted with utilization of high-quality care<sup>2</sup>. Improvements in population health will  
5 require high-quality health care over the lifespan, as global life expectancy increases and is accompanied  
6 by a growing burden of chronic non-communicable disease<sup>3,4</sup>. This is particularly true in the context of  
7 Universal Health Coverage<sup>5,6</sup>.

8  
9 The definition of “quality of care” has evolved over several decades. A seminal definition of  
10 health care quality was developed by Avedis Donabedian (1966), and is a widely-used model for  
11 measuring care quality as a function of structure, process, and outcomes<sup>7</sup>. In the 1990s, Judith Bruce and  
12 Anrudh Jain introduced a framework for assessing family planning care quality that focused on the client  
13 perspective<sup>8</sup>. Subsequent definitions and frameworks for identifying and measuring elements of quality  
14 have also been expanded and refined by the Institute of Medicine (IOM) and the World Health  
15 Organization (WHO)<sup>9</sup>, including IOM’s quality of care framework that identifies six domains of health  
16 care quality<sup>10</sup>, and WHO’s framework for quality of maternal and newborn care<sup>11</sup>. Definitions are not  
17 merely semantic; they serve as foundation for conceptual frameworks, which in turn can inform research  
18 by identifying variables of interest, formulating hypotheses for the connections between these, and  
19 devising ways to operationalize variables during data collection and analysis<sup>12, 13, 14</sup>.

20  
21 One approach to measuring care quality in global health is the Service Provision Assessment  
22 (SPA) survey, which is administered by the Demographic and Health Surveys program. SPA surveys  
23 systematically collect information about health facilities in participating countries<sup>15</sup>, are nationally  
24 representative, and have been conducted in 12 countries since 2004 (20 surveys have been conducted in  
25 total). (Prior to 2004, SPA surveys were focused on specific constructs only, such as HIV care or maternal  
26 and child health.) A key strength of SPA surveys is the use of standardized data collection tools, which

27 enables multi- and cross-country analyses<sup>16</sup>; and the comprehensive collection of information from  
28 several sources: health facility infrastructure, health workers, availability of specific health services,  
29 components of clinical care (as directly observed by data collectors), and client opinions of services  
30 received. SPA surveys have 4 overarching aims: to describe service availability; to describe readiness to  
31 provide services (infrastructure, resources, support systems); to assess whether standards of care (quality,  
32 content) are followed during service delivery; and whether clients and providers are satisfied<sup>15</sup>.

33

34 The objective of this analysis was to examine how care quality has been studied using SPA data.  
35 Measurement of quality is important for the global community, and is a key objective of SPA surveys.  
36 There have been other systematic reviews of quality of care measurement using SPA data, but these have  
37 been narrowly topic-specific (e.g., maternal and child health care or family planning only<sup>17, 18, 19</sup>). We  
38 took a broader approach in conducting a systematic review of the literature to collate all information on  
39 how SPA data have been used to study quality of care, including theoretical frameworks and definitions  
40 employed, and how these have been operationalized. Although there are a number of existing health  
41 facility assessment tools<sup>20</sup>, SPA surveys include a broad capture of data elements reflecting different  
42 aspects of care quality (including facility assessments, interviews with health workers, observations of  
43 care delivery and counseling practices, and exit interview with clients) across multiple types of health  
44 services, and are publicly available and widely used, so constituted the tool of interest here.

45

## 46 **Methods**

47 *Literature search:* Following the Preferred Reporting Items for Systematic Review and Meta-Analysis  
48 (PRISMA) guidelines, we used a keyword search for papers published since 2004 (the date of the first  
49 comprehensive, i.e. not topic-specific, SPA survey) in PubMed. (See search details in Appendix 1.) The  
50 search was conducted in January 2018. No language restriction was applied. All results were exported to  
51 Covidence software, and duplicates were removed.

52

53 *Article selection:* Each author reviewed abstracts and deemed whether the abstract should be excluded.  
54 The criterion for exclusion was if a paper clearly did not use SPA data for an empirical analysis that was  
55 published as a peer-reviewed article. The first 10% of titles and abstracts were screened by both authors,  
56 to attain inter-rater reliability; after this, the authors screened independently. Next, for non-excluded  
57 abstracts, full texts were retrieved for these publications. Each author reviewed a portion of these full  
58 texts (using the same exclusion criteria); 10% of full texts were screened by both authors before  
59 independent screening. Non-English language publications were translated using Google Translate.

60  
61 *Data extraction:* CM extracted data from the eligible full-text publications based on a pre-defined data  
62 extraction form. Covidence software was used for data extraction. Data elements extracted were: details  
63 on the SPA dataset(s) used in the article, theoretical framework applied, and how care quality was  
64 operationalized.

65  
66 *Data analysis:* Informed by the Donabedian framework, the variables used to operationalize care quality  
67 in each article were classified as relating to the constructs of Structure, Process, or Outcomes. Within  
68 Structure, 6 key domains were then identified: infrastructure; staffing; service availability; supplies,  
69 medicines, equipment; monitoring; and protocols and guides. The specific variables or data elements  
70 referenced in every included article were mapped to each of these domains, plus the domains within  
71 Process and Outcomes. These were tallied, and counts were compared across constructs and domains.

## 72 73 **Results**

74 The search strategy yielded 3250 unique abstracts to review; after screening abstracts, 2910 were  
75 excluded and 340 were selected for full-text screening (Figure 1). Ultimately 34 studies were included in  
76 this analysis (a list of all included studies is included in Appendix 2). The reasons for exclusion during  
77 full-text screening were: not being a peer-reviewed empirical paper or not including any data or empirics  
78 (n=163), including non-SPA data (n=137), and being a review article or meta-analysis (n=6).

79

80 The number of publications using SPA data to explore quality of care has increased dramatically  
81 over time (Figure 2). There is also increasing use of multiple SPA modules: health facility and provider  
82 interviews (corresponding to Structure aspects of quality), service observation (Process aspects of  
83 quality), and client exit interview (contributed to Process and Outcome aspects of quality). All but one  
84 included paper used the facility interview dataset, 21 used the provider interview dataset, 18 used at least  
85 one observation dataset, and 12 used the exit interview dataset.

86

87 The Kenya 2010 SPA dataset was most commonly used (in over half of the included papers), and  
88 approximately 30% of papers used Tanzania 2006, Namibia 2009, Rwanda 2007 or Uganda 2007 (see  
89 Appendix 3 for per-survey counts). There are some SPA surveys that did not appear in our search results  
90 (i.e., Bangladesh 2014 and Nepal 2015). Just over half (56%, n=19) of the included publications used just  
91 one SPA dataset, while the remaining 44% (n=15) used more than one; this exact pattern held even  
92 among the most recent papers (published in 2016 and 2017: 12 were single-survey papers and 10 were  
93 multi-country papers).

94

95 One-third of included papers (n=13, 38%) was informed by no theoretical or conceptual  
96 framework for care quality. Among those papers referencing a framework (n=21), the most common was  
97 the Donabedian model (used by 7 studies); an additional 12 studies used alternative existing frameworks,  
98 such as the WHO quality of care framework for maternal and newborn care<sup>11</sup>; the Primary Health Care  
99 Performance Initiative Framework (PHCPI)<sup>21</sup>, and WHO “building blocks” of health systems  
100 framework<sup>22</sup>; and 2 defined their own framework. Since the Donabedian framework was the most  
101 common in these articles and is broadly applicable (not service-specific), we use it to organize and report  
102 on quality constructs from all articles in this review.

103

104 Most papers used quality-related measures relating to structure (n=29) and/or process (n=24),  
105 while far fewer (n=11) included measures related to outcomes (Table 1). (Although SPA surveys do not  
106 capture outcomes as defined by the Donabedian framework, i.e., changes in health status, they include  
107 data on the patient experience which is how outcomes are characterized in this analysis.) Only 9 articles  
108 included variables that captured all three of these constructs; the rest mentioned only one (n=12) or two  
109 (n=13).

#### 110 Structure

111 Among the 30 papers that considered Structure as a quality construct, the most common domains  
112 were supplies/medicines, staffing, infrastructure and availability of services (Table 1). Papers commonly  
113 used the World Health Organization Service Availability and Readiness Assessment (SARA)  
114 methodology for reporting Structure elements<sup>23</sup>; the “readiness” component of SARA measures include  
115 information about inputs required for general readiness, and for providing specific types of services  
116 (availability of clinical guidelines, diagnostic tools, medicines, trained health workers).. Across all papers,  
117 30 structure variables were mentioned. At most, articles discussed 24 structure variables; the average  
118 number was 8 and the median was 6. Articles about antenatal care and primary health care incorporated  
119 (on average and median) slightly more structure variables than papers on other types of services, but there  
120 was vast heterogeneity in all service types (Appendix 4).

121



<b>Table 1: Use of domains and variables within each of the main constructs of care quality (Structure, Process, Outcomes) among articles included in this systematic review</b>		
<b>Domain</b>	<b>Variable</b>	<b>Number of studies including this variable</b>
<b>STRUCTURE (n=30)</b>		
Infrastructure	Water	n=16
	Ambulance/transport/referral	n=14
	Telephone/communication	n=13
	Electricity/light	n=13
	Infection control/waste or sharps disposal	n=13
	Soap/gloves	n=11
	Toilet/latrine	n=9
	Privacy	n=7
	Waiting area/room	n=6
	Cleanliness	n=5
	Adequate storage	n=1
Other†	n=4	
Staffing	Recently-trained staff	n=15
	Availability of trained personnel	n=10
	Number of personnel	n=5
	Years of experience	n=2
	Other‡	n=6
Service availability	Service-specific availability	n=14
	Times/days services are available	n=11
Supplies, equipment, medicines	Item availability (medicines; supplies)	n=23
	Inventory or stock ledger maintained	n=5
	Product organization: by expiration date	n=4
	Product storage: protected	n=4
	Other*	n=3
Monitoring	Recent supervisory supervision visit	n=10
	Management meetings held/system for reviewing management issues	n=10
	QA/monitoring system in place	n=9
	Client feedback system in place	n=8
	Use HMIS/other database	n=7
	Client cards used	n=2
Protocols, guides	Guidelines/protocols available/visible	n=12
	Visual/teaching aids available/used	n=6
<b>PROCESS (n=24)</b>		
	Specific clinical or care procedures implemented	n=24
	Use of recordkeeping	n=5
	Service duration (minutes)	n=4
	Use of gloves, or handwashing	n=4
	Ensured privacy or assured confidentiality	n=3
	Wait time duration (minutes)	n=2
<b>OUTCOMES (n=11)</b>		
	Client-reported problems	n=8

Client-reported satisfaction with care	n=5
Client-reported intention to return	n=2
Client-reported intention to recommend to friends/family	n=1
Other**	n=3
† Examples of "other" infrastructure: specific person/system for infrastructure repair/maintenance; functioning incinerator; number of beds ‡ Examples of "other" staffing: providers have opportunity for promotion; providers have written job description; providers have received incentives (monetary or non-monetary); providers know opportunities for promotion * Examples of "other" supplies and equipment: proper final sterilization process used for medical equipment; medication stocking frequency; staff knowledge of processing time for equipment ** Examples of "other" outcomes: correct use of treatment; client knowledge after care	

123

124

AUTHOR FINAL VERSION

125

126 Certain aspects of infrastructure were mentioned much more commonly than others. More than  
127 two-thirds of studies reporting on infrastructure discussed availability of water, infection control/safe  
128 disposal measures, and an ambulance; while fewer than 40% discussed cleanliness, privacy for clients, or  
129 availability of a client waiting area.

130 Within supplies/medicines, the presence of specific items was the most common representation  
131 (availability of equipment and/or consumables [medicines, vaccines] was mentioned by 23 articles), but  
132 only 4 articles mentioned whether the products were well-stored (organized by expiration date) and 4  
133 mentioned whether storage was protective from the elements or damage. Within the same service type,  
134 articles analyzed availability of very different supplies and medicines (Appendix Table 4); for example,  
135 only two articles on child health care included data about medicines availability: one only assessed  
136 availability of antibiotics, and the other included medicines ranging from antibiotics to oral rehydration  
137 salts to deworming tablets. Among articles about quality of antenatal care (n=10), all but 3 included  
138 availability of iron and/or folic acid tablets, but only 4 included antimalarial medications.

139 Training of staff was much more commonly mentioned than availability of staff (e.g., round-the-  
140 clock clinician availability) or number of personnel overall. Most articles that discussed monitoring  
141 mentioned supportive supervision visits (n=10) and systems for reviewing management issues (including  
142 management meetings) (n=10), but fewer mentioned approaches for eliciting client feedback (n=9) or use  
143 of HMIS or other data systems (n=7).

144

#### 145 Process

146 There were 24 papers that included Process in their operationalization of care quality (Table 1).  
147 All used variables that captured clinical care components; these were specific to the type of visit (family  
148 planning consultation, antenatal care, childbirth, or sick child care). Different articles, even about the  
149 same type of health service, reported on process to a very different extent and using different indicators.

150 In each category of care type, between 20-33% of included articles mentioned 2 or fewer care activities.  
 151 See Appendix tables 5-8 for details on all activities included, by service type and per article.

152 *Antenatal care:* There were 10 articles that used SPA data to investigate quality of antenatal care services,  
 153 and 9 of them included variables on specific care activities. Across the included articles, 26 different care  
 154 activities were identified. The most commonly-reported activities were testing for anemia (n=8), testing  
 155 for urine protein (n=8), testing for STIs (n=6), measuring blood pressure (n=7), and measuring weight  
 156 (n=6). Only two articles included all 5 of these common activities. Many components of the WHO  
 157 essential practice guidelines for antenatal care<sup>24</sup> – including discussing pregnancy history, asking about  
 158 danger signs, examining for signs of anemia, and counseling on family planning – were reported by 3 or  
 159 fewer articles.

160 *Childbirth:* Among the 12 articles that analyzed quality of childbirth care using SPA data, 11 included  
 161 information about specific care activities; these 11 articles identified a total of 24 unique childbirth care  
 162 activities. Encouraging immediate breastfeeding (n=7), conducting uterine massage after delivery of the  
 163 placenta (n=6), ensuring skin-to-skin contact after birth (n=6), and giving oxytocin within one minute of  
 164 delivery (n=6) were the most frequently-assessed activities, but only two articles included all 4 of these  
 165 commonly-reported activities. Again, many of the WHO essential practices for childbirth<sup>24</sup> were largely  
 166 omitted, such as giving the baby vitamin K (n=2), providing antibiotic eye drops or ointment (n=2), and  
 167 conducting a vaginal exam (n=3).

168 *Pediatrics:* There were 6 articles about sick child care, and 5 of these included information about specific  
 169 care processes. In total, these 5 articles identified 24 specific care activities. The most frequently reported  
 170 activities were taking child's temperature (n=3), and asking about child's vaccination history (n=3). Few  
 171 papers included WHO's Integrated Management of Childhood Illness (IMCI)-recommended counseling  
 172 activities, such as counseling on feeding habits and illness symptoms (n=2), or informing the caretaker of  
 173 the child's diagnosis (n=2). Most activities were only included once across the five studies (felt for fever,  
 174 checked for dehydration, asked about child's Vitamin A status); and only 1 paper reported the majority of  
 175 these activities (n=20).

176 *Family planning*: The 9 articles about family planning care quality included 8 that analyzed process; and  
177 these 8 articles identified a total of 25 specific family planning care activities. The most commonly  
178 reported activities were taking client's reproductive history, blood pressure and weight (n=5), asking  
179 about chronic illnesses, smoking habits and STI symptoms (n=5), counseling on proper use of family  
180 planning methods (n=4), and ensuring client privacy (n=4). Only 2 papers included activities assessing  
181 provider adherence for clinical procedures (for injectable contraceptives or implants). Other  
182 recommended counseling practices from the Global Handbook for Family Planning Providers, such as  
183 counseling on side effects and/or when to return (n=3), partner's attitude towards family planning (n=2),  
184 and method protection against STIs/HIV (n=2) were largely omitted.

185 Additionally, 3 papers touched on issues related to privacy during the visit. Four papers included  
186 a variable related to service interaction duration, and 3 looked at wait time. Four looked at hygiene during  
187 the visit (provider handwashing and/or wearing gloves) and 5 included a measure of whether the provider  
188 referenced the client card or a facility register.

189

#### 190 Outcomes

191 Only 11 papers included such outcomes when measuring care quality (Table 1). The most  
192 common experiential outcome was whether the client reported specific problems relating to services that  
193 day (n=8), and whether the client stated they were satisfied with care received (n=5). A smaller number of  
194 articles included whether the client said they intended to return for services, and whether they would  
195 recommend the care to family or friends. Other outcomes measures included the patient's knowledge after  
196 the visit, and whether correct treatment was administered.

197

#### 198 Summarizing data elements

199 There was also substantial variation in methods used to create summary quality measures (Appendix 9).  
200 One-fifth of articles (n=7, 20%) used principal component analysis (PCA) for quantifying quality; among  
201 these, some included only structure constructs in the PCA (n=3), some included structure and process

202 measures (n=2), and some used client-reported outcomes only (n=2). Four articles calculated means  
203 within domains, and used the average of these means as an overall summary measure of quality. All other  
204 articles used simple additive indices of constructs (using binary or categorical variables).

205

206

## 207 **Discussion**

208 Publicly and freely available, SPA program data enable examination of health care quality in low-  
209 and middle-income countries, and, through the use of standardized questionnaires, cross-country analyses  
210 and comparisons. Accordingly, publications based on analyses of SPA data are increasing dramatically,  
211 and nearly half of these papers incorporate more than one SPA dataset.

212

213 The results of this review indicate opportunities for strengthening the use of SPA data for  
214 measuring care quality. Many papers do not use a theoretical framework, despite a rich literature on  
215 theories of quality and numerous service-specific frameworks developed by agencies like the World  
216 Health Organization. The use of a theoretical framework is advisable from the perspective of research  
217 rigor, reproducibility and generalizability; and also since purely empirical approaches for operationalizing  
218 care quality (e.g., principal component analysis if it is not informed by an underlying theoretical  
219 framework) result in scores that are highly sensitive to the exact method deployed<sup>19</sup> or the health system  
220 context<sup>25</sup>, which limit generalizability of the results.

221

222 We found that certain SPA modules are relatively rarely utilized – particularly service  
223 observation and client exit interviews – and as a result, certain quality constructs (especially patient  
224 reports of the care experience, which were characterized here as “outcomes”) are often omitted from  
225 characterizations of care quality. It should be noted that patient satisfaction can be challenging to measure  
226 with high validity and reliability<sup>26, 27, 28</sup>, and reported experience may also depend on patients’ overall

227 expectation of the health system and care quality<sup>29,30,31</sup>. Future quality of care studies – including SPA  
228 surveys – should explore ways to improve measurement of client-reported experience.

229  
230 The vast heterogeneity in operationalization limits the development of a larger literature on care  
231 quality, since these studies differ in their underlying approach to defining and measuring quality. A multi-  
232 dimensional approach to measuring health care quality may be necessary due to the complexity and  
233 multifaceted nature of the interaction<sup>32</sup> – for example, studies have found a low correlation between  
234 “quality” when measured as inputs (infrastructure) versus service delivery<sup>33</sup>, as well as a weak  
235 association between service quality and patients’ satisfaction with care<sup>34</sup>.

236  
237  
238 These results add to a growing literature about opportunities to expand our learning from public  
239 goods datasets like SPA. Recent reviews have identified gaps between global standards of care quality,  
240 and studies using publicly-available data (including SPA)<sup>17, 18, 35</sup>. Although disagreement and debate  
241 about how to define quality is not new, nor unique to SPA data, the use of a common dataset with  
242 standardized tools offers a unique opportunity to align definitions and methods. In Table 2, we suggest  
243 ways in which care quality can more comprehensively be measured using available SPA data elements,  
244 based on gaps identified in this study. First, structural aspects of care quality should include infrastructure  
245 relevant to the patient experience<sup>36,37</sup>. Second, many analyses use the SARA methodology for capturing  
246 structure and process; we recommend the development of a new harmonized “SARA+” approach, to  
247 represent quality of the components themselves. For example, SARA measures whether essential  
248 medicines are available, but a SARA+ metric would include whether these medicines are stored  
249 appropriately; SARA includes recent training of personnel, and SARA+ would capture the average  
250 availability of these personnel. Third, quality measures using SPA data should strive to include outcomes  
251 – whether patient-reported or objectively measured<sup>38</sup>. SPA surveys are a unique source of information  
252 from service observations and exit interviews, and therefore include rare information about outcomes

253 related to patient experience, which would enhance our understanding of care quality. We also  
254 recommend that future SPA surveys expand the exit interview module, to collect more nuanced and rich  
255 data about the care experience; and add more service types to the direct observation modules, beyond the  
256 maternal and child health care services currently included. There are challenges in designing approaches  
257 to measuring and reporting care quality, even with standardized datasets, particularly across countries –  
258 certain constructs may only be relevant in certain settings due to disease endemicity, for example. Future  
259 efforts to standardize quality measurement using SPA data should consider how to account for such  
260 geographic heterogeneity and recommend ways in which multi-country analyses might overcome this  
261 challenge.

262  
263         Some limitations to these findings should be noted. First, the search was only conducted in  
264 PubMed so we may have missed papers indexed in other databases. Second, the SPA program is only one  
265 source of standardized health service quality data; SPA data were selected for this study do to their  
266 comprehensiveness and the inclusion of multiple data sources and perspectives (facility, provider,  
267 patient), but other datasets, including from the WHO and the World Bank, should also be included for a  
268 more comprehensive understanding of the state of quality measurement in low- and middle-income  
269 countries. Third, the classification of variables into quality constructs was based on our assessment (not  
270 necessarily how the authors categorized these) which was a subjective process. We chose to use WHO  
271 service guidelines for classifying process measures, but other frameworks might also be informative<sup>39, 40</sup>.  
272 Fourth, true “outcomes” are also not included in SPA surveys, so this analysis could only include  
273 outcomes insofar as they related to patient-reported experience. (SPA surveys would be even more  
274 informative and powerful if they included information on true “outcomes” such as morbidity and  
275 mortality (whether through the SPA data collection process, or through linkages with health information  
276 systems). Lastly, this review did not assess quality of the included studies, nor compare results of these  
277 quality assessments.

278



279 **Conclusion**

280 Consistency in quality measurement is critical for cross-country analyses and comparisons.  
281 Future studies should seek to refine and harmonize these measures, and to evaluate the advantages and  
282 disadvantages of methods to summarize across Structure, Process and Outcome domains, including  
283 comparing different empirical approaches for summarizing quality of care measures. There may also be  
284 important distinctions across service types (e.g. sick child care, antenatal care, childbirth or family  
285 planning services), and this area also merits further study. SPA surveys are an essential tool for  
286 researchers and policymakers, and the public health community should strive to maximize the impact and  
287 global learnings from studies that use these data by developing consensus around definitions, frameworks  
288 and operationalizations of key variables.

289

AUTHOR FINAL VERSION

Table 2: Suggestions for strengthening and harmonizing measurement of care quality using SPA data

- Incorporate patient-facing elements of infrastructure, e.g. availability of toilet, adequate waiting area, auditory and visual privacy.
- Enhance SARA indicators by adding quality dimensions, e.g. storage quality and organization, capacity to repair equipment, employment quality (opportunities for promotion, job description), hours and days of service/staff availability.
- Include outcomes including patient-reported (“complaint score” of problems during visit, satisfaction with care, would recommend to others) and appropriateness of clinical care.

290

AUTHOR FINAL VERSION

291 **Co-Author Contact Information:**

292 Corrina Moucheraud: University of California Los Angeles, [cmoucheraud@ucla.edu](mailto:cmoucheraud@ucla.edu)

293 Kaitlyn McBride: University of California Los Angeles, [kaitlynbmcbride@gmail.com](mailto:kaitlynbmcbride@gmail.com)

294

295 **Acknowledgments:** The authors thank Bethany Myers of the UCLA Biomedical Library for her expert  
296 guidance on refining the search strategy and conducting the search.

297

298 **Financial Support:** The authors benefitted from facilities and resources provided by the California  
299 Center for Population Research at UCLA (CCPR), which receives core support (P2C-HD041022) from  
300 the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). KM  
301 receives training support from CCPR, funded by the Eunice Kennedy Shriver National Institute of Child  
302 Health and Human Development (T32-HD007545). CM receives support from the UCLA Clinical and  
303 Translational Science Institute, funded by the NIH/National Center for Advancing Translational Sciences  
304 (KL2TR001882).

305

306 **Disclosures:** The authors declare no conflicts of interest.

307

308 **Author current addresses:**

309 Corrina Moucheraud: University of California Los Angeles, Department of Health Policy and

310 Management, 650 Charles E. Young Drive S., Los Angeles CA 90095 USA; [cmoucheraud@ucla.edu](mailto:cmoucheraud@ucla.edu)

311 Kaitlyn McBride: University of California Los Angeles, Department of Health Policy and Management,

312 650 Charles E. Young Drive S., Los Angeles CA 90095 USA; University of California Los Angeles,

313 [kaitlynbmcbride@gmail.com](mailto:kaitlynbmcbride@gmail.com)

314

315 **References**

- 316 1. World Health Organization, Organisation for Economic Co-operation and Development, The  
 317 World Bank, 2018. Delivering quality health services: A global imperative for universal health  
 318 coverage. Geneva.
- 319 2. Kruk ME, Gage AD, Joseph NT, Danaei G, García-Saisó S, Salomon JA, 2018. Mortality due to low-  
 320 quality health systems in the universal health coverage era: A systematic analysis of amenable  
 321 deaths in 137 countries. *The Lancet*.
- 322 3. Hum RJ, Verguet S, Cheng Y-L, McGahan AM, Jha P, 2015. Are global and regional improvements  
 323 in life expectancy and in child, adult and senior survival slowing? *PloS One* 10: e0124479.
- 324 4. Atun R, Jaffar S, Nishtar S, Knaul FM, Barreto ML, Nyirenda M, Banatvala N, Piot P, 2013.  
 325 Improving responsiveness of health systems to non-communicable diseases. *The Lancet* 381:  
 326 690-697.
- 327 5. National Academies of Sciences E, Medicine, 2018. Crossing the Global Quality Chasm:  
 328 Improving Health Care Worldwide. Washington, DC: The National Academies Press.
- 329 6. Boerma T, Eozenou P, Evans D, Evans T, Kieny M-P, Wagstaff A, 2014. Monitoring progress  
 330 towards universal health coverage at country and global levels. *PLoS Medicine* 11: e1001731.
- 331 7. Donabedian A, 1988. The quality of care: How can it be assessed? *JAMA* 260: 1743-1748.
- 332 8. Bruce J, 1990. Fundamental elements of the quality of care: A simple framework. *Studies in*  
 333 *Family Planning* 21: 61-91.
- 334 9. World Health Organization, 2015. WHO global strategy on people-centred and integrated health  
 335 services. Geneva.
- 336 10. Baker A, 2001. Crossing the quality chasm: a new health system for the 21st century: British  
 337 Medical Journal Publishing Group, 1192.
- 338 11. Tunçalp Ö, Were W, MacLennan C, Oladapo O, Gülmezoglu A, Bahl R, Daelmans B, Mathai M,  
 339 Say L, Kristensen F, 2015. Quality of care for pregnant women and newborns—the WHO vision.  
 340 *BJOG: an international journal of obstetrics & gynaecology* 122: 1045-1049.
- 341 12. Aneshensel CS, 2012. Theory-based data analysis for the social sciences: Sage Publications.
- 342 13. McGlynn EA, 1997. Six challenges in measuring the quality of health care. *Health Affairs* 16: 7-  
 343 21.
- 344 14. King G, Keohane RO, Verba S, 1994. Designing social inquiry: Scientific inference in qualitative  
 345 research: Princeton University Press.
- 346 15. Demographic and Health Surveys Program, n/d. SPA Overview. Available at:  
 347 <http://www.dhsprogram.com/What-We-Do/Survey-Types/SPA.cfm>. Accessed.
- 348 16. Alva S, Kleinau E, Pomeroy A, Rowan K, 2009. Measuring the impact of health systems  
 349 strengthening: a review of the literature. Washington: US Agency for International  
 350 Development.
- 351 17. Sheffel A, Karp C, Creanga AA, 2018. Use of Service Provision Assessments and Service  
 352 Availability and Readiness Assessments for monitoring quality of maternal and newborn health  
 353 services in low-income and middle-income countries. *BMJ Global Health* 3: e001011.
- 354 18. Brizuela V, Leslie HH, Sharma J, Langer A, Tunçalp Ö, 2019. Measuring quality of care for all  
 355 women and newborns: how do we know if we are doing it right? A review of facility assessment  
 356 tools. *The Lancet Global Health*.
- 357 19. Mallick L, Temsah G, Wang W, 2019. Comparing summary measures of quality of care for family  
 358 planning in Haiti, Malawi, and Tanzania. *PloS One* 14: e0217547.
- 359 20. Nickerson JW, Adams O, Attaran A, Hatcher-Roberts J, Tugwell P, 2014. Monitoring the ability to  
 360 deliver care in low-and middle-income countries: a systematic review of health facility  
 361 assessment tools. *Health policy and planning* 30: 675-686.

- 362 21. Veillard J, Cowling K, Bitton A, Ratcliffe H, Kimball M, Barkley S, Mercereau L, Wong E, Taylor C,  
 363 Hirschhorn LR, 2017. Better measurement for performance improvement in Low-and Middle-  
 364 Income countries: the primary health care performance Initiative (PHCPI) experience of  
 365 conceptual framework development and indicator selection. *The Milbank Quarterly* 95: 836-  
 366 883.
- 367 22. Organization WH, 2007. Everybody's business--strengthening health systems to improve health  
 368 outcomes: WHO's framework for action.
- 369 23. Health Statistics and Information Systems, 2015. Service Availability and Readiness Assessment  
 370 (SARA): Reference manual v2.2. World Health Organization, ed.
- 371 24. Organization WH, Health WHOR, Health WHODOR, UNICEF., Activities UNFFP, 2003. Pregnancy,  
 372 childbirth, postpartum, and newborn care: a guide for essential practice: World Health  
 373 Organization.
- 374 25. Jackson EF, Siddiqui A, Gutierrez H, Kanté AM, Austin J, Phillips JF, 2015. Estimation of indices of  
 375 health service readiness with a principal component analysis of the Tanzania Service Provision  
 376 Assessment Survey. *BMC Health Services Research* 15: 536.
- 377 26. Dunsch F, Evans DK, Macis M, Wang Q, 2018. Bias in patient satisfaction surveys: a threat to  
 378 measuring healthcare quality. *BMJ global health* 3: e000694.
- 379 27. Voutilainen A, Pitkääho T, Kvist T, Vehviläinen-Julkunen K, 2016. How to ask about patient  
 380 satisfaction? The visual analogue scale is less vulnerable to confounding factors and ceiling  
 381 effect than a symmetric Likert scale. *Journal of advanced nursing* 72: 946-957.
- 382 28. Ross CK, Steward CA, Sinacore JM, 1995. A comparative study of seven measures of patient  
 383 satisfaction. *Medical care*.
- 384 29. Roder-DeWan S, Gage AD, Hirschhorn LR, Twum-Danso NAY, Liljestrand J, Asante-Shongwe K,  
 385 Rodríguez V, Yahya T, Kruk ME, 2019. Expectations of healthcare quality: A cross-sectional study  
 386 of internet users in 12 low- and middle-income countries. *PLOS Medicine* 16: e1002879.
- 387 30. Bleich SN, Özaltın E, Murray CJ, 2009. How does satisfaction with the health-care system relate  
 388 to patient experience? *Bulletin of the World Health Organization* 87: 271-278.
- 389 31. Bowling A, Rowe G, McKee M, 2013. Patients' experiences of their healthcare in relation to their  
 390 expectations and satisfaction: a population survey. *Journal of the Royal Society of Medicine* 106:  
 391 143-149.
- 392 32. Hanefeld J, Powell-Jackson T, Balabanova D, 2017. Understanding and measuring quality of care:  
 393 Dealing with complexity. *Bulletin of the World Health Organization* 95: 368.
- 394 33. Leslie HH, Sun Z, Kruk ME, 2017. Association between infrastructure and observed quality of  
 395 care in 4 healthcare services: A cross-sectional study of 4,300 facilities in 8 countries. *PLoS*  
 396 *Medicine* 14: e1002464.
- 397 34. Diamond-Smith N, Sudhinaraset M, Montagu D, 2016. Clinical and perceived quality of care for  
 398 maternal, neonatal and antenatal care in Kenya and Namibia: The service provision assessment.  
 399 *Reproductive Health* 13: 92.
- 400 35. Moxon SG, Guenther T, Gabrysch S, Enweronu-Laryea C, Ram PK, Niermeyer S, Kerber K, Tann  
 401 CJ, Russell N, Kak L, 2018. Service readiness for inpatient care of small and sick newborns: What  
 402 do we need and what can we measure now? *Journal of Global Health* 8.
- 403 36. Coulter A, 2017. Measuring what matters to patients. *BMJ* 356: j816.
- 404 37. Mohammed K, Nolan MB, Rajjo T, Shah ND, Prokop LJ, Varkey P, Murad MH, 2016. Creating a  
 405 patient-centered health care delivery system: a systematic review of health care quality from  
 406 the patient perspective. *American Journal of Medical Quality* 31: 12-21.
- 407 38. Browne K, Roseman D, Shaller D, Edgman-Levitan S, 2010. Measuring patient experience as a  
 408 strategy for improving primary care. *Health Affairs* 29.

- 409 39. Tripathi V, Stanton C, Strobino D, Bartlett L, 2015. Development and validation of an index to  
410 measure the quality of facility-based labor and delivery care processes in sub-Saharan Africa.  
411 PloS one 10.
- 412 40. Spector JM, Lashoher A, Agrawal P, Lemer C, Dziekan G, Bahl R, Mathai M, Merialdi M, Berry W,  
413 Gawande AA, 2013. Designing the WHO Safe Childbirth Checklist program to improve quality of  
414 care at childbirth. International Journal of Gynecology & Obstetrics 122: 164-168.
- 415

AUTHOR FINAL VERSION