Satisfaction measurement instruments for healthcare service users: a systematic review

Instrumentos para mensuração de satisfação de usuários de serviços de saúde: uma revisão sistemática

Los instrumentos para medir la satisfacción del usuario de los servicios de salud: una revisión sistemática

Abstract

Patient satisfaction surveys can be an interesting way to improve quality and discuss the concept of patient-centered care. This study aimed to conduct a systematic review of the validated patient satisfaction measurement instruments applied in healthcare. The systematic review searched the MEDLINE/PubMed, LILACS, SciELO, Scopus and Web of Knowledge. The search strategy used the terms: “Patient Satisfaction” AND “Patient centered care” AND “Healthcare survey OR Satisfaction questionnaire” AND “Psychometric properties”. 37 studies were included and almost all studies showed that satisfaction is a multidimensional construct. In these studies, 34 different instruments were used and most surveys contained the dimension patient-healthcare professional interactions, physical environment and management process. The COSMIN score for methodological quality showed that most of them scored a good or fair average. We can conclude that there is not a gold standard instrument for patient satisfaction assessment but some dimensions are essential for this construct.

Quality of Health Care; Patient-Centered Care; Patient Satisfaction; Psychometrics

Resumo


Qualidade da Assistência à Saúde; Assistência Centrada no Paciente; Satisfação do Paciente; Psicometria
Background

In the past decade, several programs for quality improvement were developed in many health services and systems. Leatherman et al. reported that there are publications on this subject in various regions of the world, including India, South America, Asia and North America. More specifically, on the topic of how to increase quality, many institutions have been trying to establish guidelines and parameters for quality assessment and improvement.

Patient-centered care has been accepted as one of the pillars for health services to achieve quality of care. This issue must involve the patients’ needs and preferences in decision-making about care, as well as previous experience within these services. Thus, the literature shows a trend towards studying what actually is relevant for the patient’s experience within healthcare.

Considering this goal, patient satisfaction studies can be a useful way of measuring a part of this complex construct, because it is an important concern related to quality of healthcare and patient centeredness, both in the public and private sectors. The concept of satisfaction is well known and applied in management and especially in marketing. Evrard defines satisfaction as a psychological state, appearing after consumption (of a service) and compared to an initial baseline (services’ expectation). Olivier and Fournier & Mick expressed satisfaction as the subjective comparison between expectations and perceptions of service performance. The satisfaction is subjective, yet measurable, provided that the consumer has an initial point of reference. Satisfaction on healthcare must consider the “emotional” aspect related to a medical action that may have a significant impact on the patients’ own health. In this way, understanding all dimensions and factors related to patient satisfaction is important in order to improve the quality of healthcare.

Recent publications on quality of care have discussed the challenge of improving quality and most of them have stressed patient-centered care as one of the core goals. In this context, the user satisfaction construct has already been scrutinized by many different study designs in different areas. The validation and cross-cultural adaptation of surveys are the most common study designs, but, up to now, this area has needed more investigation.

Choosing the appropriate instrument for a patient-related outcome measurement is difficult; it depends on the proposed use and can probably have an impact on the accuracy of the results. The discussion about which are the essential psychometric properties for healthcare instruments validation is still lacking in the literature. It is therefore quite important to know if there is some reference instrument to measure patient satisfaction with healthcare services and if there is a similar scope in the structure of these instruments after the validation process. Given this concern, this study aims to perform a systematic review of validation studies of patient’s satisfaction measurement instruments applied in healthcare services. It thus will seek to identify the kinds of tools and psychometric properties analyzed and the quality of the instruments.

Methods

Study design and information sources

A systematic review of patient satisfaction surveys was undertaken. Electronic databases were chosen based on their indexed journals profile. MEDLINE/PubMed, LILACS (Virtual Health Library; Bireme/PAHO), SciELO (Scientific Electronic Library Online), Scopus and Web of Knowledge were used as reference material in this study. LILACS and SciELO were included due to the fact that both have Latin American studies in their collections. All databases were screened in December, 2013. The search strategy was oriented only for validation studies of patient satisfaction instruments and for this reason other terms that could be relevant for quality of care or other pillars of this complex construct were not included. In this way, terms that could have other concepts inputted were not included, such as patient preference, patient needs and consumer satisfaction. In addition, the term “validation study” was not used, because it was only introduced as a topic in MEDLINE in 2008. After testing several words and MeSH terms, the search strategy used the terms: “Patient Satisfaction[MeSH Terms]” AND “Patient centered care[MeSH Terms]” AND “Healthcare survey OR Satisfaction questionnaire” AND “Psychometric properties”.

Inclusion and exclusion criteria

The research assumed the following eligibility criteria: (a) only studies published in English, Spanish, French and Portuguese; (b) studies published between 2002 and 2013; (c) studies which have used a patient satisfaction tool to investigate only the opinion of health service users; (d) studies that had evaluated and described at least one step of the COSMIN checklist were considered. The temporal limitation was justified in the
report Crossing the Quality Chasm: A New Health System for the 21st Century 3, published in 2001, where patient-centered care as a dimension of quality of care has become more widespread and highlighted a renewed place for the patient satisfaction. No types of health service were excluded, in order to guarantee different patient profiles. The definition of health service used was “any center that provides diagnosis, treatment or prevention for more than one kind of disease, illness or injury in the human body”. However, instruments used to measure patient satisfaction regarding just one specific treatment or illness or clinical condition were excluded, such as diabetes or chemotherapy, for instance. Systematic review and metanalysis studies were excluded.

**Study selection**

Two independent reviewers applied the inclusion and exclusion criteria to the selected articles. All titles and abstracts were screened after the removal of duplicates and based on this examination, a list of references for full text reading was prepared. After this, both reviewers selected the papers to be included in the next step (qualitative synthesis of the studies). The search results and papers screened were discussed by e-mail and also during some meetings. Any disagreements on the records included in this review were submitted to a third reviewer.

**Data extraction**

A form was prepared to collect priority information; the following items were extracted from each study: year, authors, study design (instrument validation, adaptation – changes suggested from the original, and cross-cultural adaptation), healthcare setting (hospital, clinic), satisfaction instrument utilized, dimensions of the instrument and psychometric properties analyzed.

**Evaluation of methodological quality**

It is necessary to emphasize that the methodological quality of the study is as important as the quality of a measurement instrument. The psychometric properties of an instrument are merely one part of the validation process and, thus, other steps must be demonstrated during this process 14,15. The COSMIN checklist is one of the guidelines to evaluate the methodological quality of measurement instruments in different areas, and many systematic reviews have already proved its usefulness 17,18,19. Considering this, the present study chose the COSMIN checklist to evaluate the methodological quality of the studies included in the reviewing process. This checklist is a standardized tool to assess studies on measurement properties. It contains nine boxes and each one has a different measurement property, with 5-18 items per box on design aspects and statistical methods. The boxes refer to: internal consistency, reliability, measurement error, content validity, structural validity, hypotheses testing, cross-cultural validity, criterion validity, and responsiveness 17.

Each measurement property is evaluated separately. This means that if multiple measurement properties are assessed in one study, several COSMIN boxes need to be completed. The content evaluated in each box can be judged as: “excellent”; “good”; “fair”; “poor”. The methodological quality score per box is obtained by taking the lowest rating of any item in the box (“worse score counts”) 17. The checklist includes one additional box for studies that used item response theory (IRT). Two additional boxes are recommended (interpretability and generalizability of results), but, as they do not provide a score, they were not used here. Here the COSMIN checklist was applied for each selected study by the principal reviewer and all doubts were discussed.

**Results**

The initial search strategy found 1,398 articles, and after exclusion of duplicates, the total was 1,286. After the inclusion and exclusion criteria were applied, there were only 49 articles remaining (Figure 1). All these studies were retrieved for reading and 12 were excluded in this phase. By the end of this phase, 37 articles were included in the qualitative synthesis of the reviewing process (Figure 1).

**General data**

The results showed that satisfaction measurement instruments have been applied all around the world (Europe, Asia, North America, Australia, and Africa). The sample size in the studies screened showed a wide variation. The number of people included ranged from 80 20 to 7,093 21; almost all studies collected some kind of cultural and socio demographic variables of the patients (age, gender, family income, educational level and health area of treatment).

It was difficult to identify which health services were public or private, most studies did not clearly state this 20,21. Only five studies 22,23,24,25,26 clearly stated the ownership of the service studied. Service complexity level showed major differences because three studies utilized university...
hospitals to measure patient satisfaction, 10 used only clinics and another 20 studies measured patient satisfaction in general hospitals (Table 1). There are many different instruments being used, whether in hospitals or in clinics. 34 different instruments for satisfaction measurement were found and few of them were utilized by more than one study: Consumer Emergency Care Satisfaction Scale (CECSS) 26,27 – twice; Echelle de Qualité des Soins en Hospitalisation (EQS-H scale) 28,29 – twice; Primary Care Satisfaction Survey for Women (PCSSW) 30,31 – twice (Table 1). It seems that there is not a gold standard or a standardized instrument for satisfaction measurements in health services.

Methodological quality

The most frequent study design observed was instrument validation and only six cross-cultural adaptations were found 26,30,32,33,34,35 (Table 1).

An analysis of the psychometric properties indicates that most studies used the classical test theory (CTT) to evaluate the instruments’ structure. Only one paper used the IRT 19. The main properties assessed were structural validity and internal consistency. Just five papers were found not to approach these two proprieties together 18,36,37,38,39. Content validity (face validity), hypotheses testing, reliability (test-retest) and cross-cultural validity were other properties analyzed in many studies (Table 1).

The COSMIN checklist scores demonstrated that most papers achieved a medium score (“good” or “fair”) on the methodology quality of the psychometric analysis. Only four studies demonstrated low scores (“poor”) 23,37,40,41 and eleven studies had at least one property categorized as excellent 21,24,31,36,39,42,43,45,46,47,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56 presented the Cronbach α value as an internal consistency measure (ranged from 0.39 to 0.99); most of them showed correlations between items, subscales and other satisfaction measures (Table 1). The main statistical methods utilized were the exploratory factorial analysis (EFA) and confirmatory factorial analysis (CFA) to evaluate the instruments’ structure.
Table 1

Characteristics of the studies.

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>N</th>
<th>Health care setting</th>
<th>Study design</th>
<th>Instrument used</th>
<th>Construct dimensions (*)</th>
<th>Measurement properties **</th>
<th>Psychometric values</th>
<th>COSMIN scores **</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meehan et al. 45 (2002)</td>
<td>356</td>
<td>Hospital validation</td>
<td>Inpatient Evaluation of Service Questionnaire (IESQ)</td>
<td>Multidimensional (3)</td>
<td>1-Internal consistency</td>
<td>a &gt; 0.86;</td>
<td>1-Fair</td>
<td></td>
</tr>
<tr>
<td>Hendriks et al. 50 (2004)</td>
<td>728</td>
<td>Hospital adaptation</td>
<td>Satisfaction with Hospital Care Questionnaire</td>
<td>Multidimensional (12)</td>
<td>1-Internal consistency</td>
<td>a &gt; 0.6;</td>
<td>1-Fair</td>
<td></td>
</tr>
<tr>
<td>Scholle et al. 31 (2004)</td>
<td>1,202</td>
<td>Clinic validation</td>
<td>Primary Care Satisfaction Survey for Women (PCSSW)</td>
<td>Multidimensional (3)</td>
<td>1-Structural validity</td>
<td>a &gt; 0.80;</td>
<td>1-Good</td>
<td></td>
</tr>
<tr>
<td>Lam et al. 34 (2005)</td>
<td>711</td>
<td>Hospital trans cultural adaptation</td>
<td>Medical interview Satisfaction Scale (C-MISS)</td>
<td>Bidimensional</td>
<td>1-Internal consistency</td>
<td>a &gt; 0.84;</td>
<td>1-Good</td>
<td></td>
</tr>
<tr>
<td>Gonzalez et al. 42 (2005)</td>
<td>1,910</td>
<td>Hospital validation</td>
<td>In-patient Satisfaction Questionnaire</td>
<td>Multidimensional (6)</td>
<td>1-Internal consistency</td>
<td>a &gt; 0.60;</td>
<td>1-Excellent</td>
<td></td>
</tr>
<tr>
<td>Wilde &amp; Hendriks 36 (2005)</td>
<td>262</td>
<td>Clinic adaptation</td>
<td>Client Satisfaction Questionnaire (CSQ)</td>
<td>Unidimensional</td>
<td>1-Structural validity</td>
<td>a = 0.92;</td>
<td>1-Good</td>
<td></td>
</tr>
<tr>
<td>Gerkensmeyer &amp; Austin 22 (2005)</td>
<td>278</td>
<td>Clinic Children Mental Health Service validation</td>
<td>Parent Satisfaction Scale (PSS)</td>
<td>Unidimensional</td>
<td>1-Internal consistency</td>
<td>a &gt; 0.70;</td>
<td>1-Fair</td>
<td></td>
</tr>
<tr>
<td>Otani et al. 53 (2005)</td>
<td>8,465</td>
<td>Hospital adaptation</td>
<td>Patient Satisfaction Survey</td>
<td>Multidimensional (3)</td>
<td>1-Internal consistency</td>
<td>a &gt; 0.8;</td>
<td>1-Fair</td>
<td></td>
</tr>
</tbody>
</table>

(continues)
<table>
<thead>
<tr>
<th>Study (year)</th>
<th>N</th>
<th>Health care setting</th>
<th>Study design</th>
<th>Instrument used</th>
<th>Construct dimensions (*)</th>
<th>Measurement properties **</th>
<th>Psychometric values</th>
<th>COSMIN scores **</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poinsot et al. 40 (2006)</td>
<td>416</td>
<td>Hospital (Oncology Center)</td>
<td>Validation</td>
<td>Out-Patsat3S</td>
<td>Bidimensional</td>
<td>1-Structural validity 2-Internal consistency</td>
<td>α &gt; 0.90; inter-item correlations (r = 0.33-0.91); correlation with other instrument (r = 0.38)</td>
<td>1-Poor 2-Poor</td>
</tr>
<tr>
<td>Tso et al. 25 (2006)</td>
<td>344</td>
<td>Hospital</td>
<td>Validation</td>
<td>Concise Outpatient Department User Satisfaction Scale</td>
<td>Bidimensional</td>
<td>1-Internal consistency 2-Structural validity 3-Hypotheses testing</td>
<td>α &gt; 0.90; Guttman split-half reliability (0.93); Spearman-Brown reliability (0.94)</td>
<td>1-Fair 2-Fair 3-Fair</td>
</tr>
<tr>
<td>Gagnon et al. 47 (2006)</td>
<td>873</td>
<td>Clinic</td>
<td>Validation</td>
<td>Health Care Satisfaction Questionnaire (HCSQ)</td>
<td>Multidimensional (3)</td>
<td>1-Structural validity 2-Internal consistency 3-Content validity 4-Reliability</td>
<td>α &gt; 0.74; KMO (0.95); CFI = 0.93; NFI = 88; AGFI = 0.87; RMSEA = 0.6; correlations between factor (r &gt; 0.6); ICC (&gt; 0.45)</td>
<td>1-Excellent 2-Excellent 3-Excellent 4-Fair</td>
</tr>
<tr>
<td>Campbell et al. 49 (2007)</td>
<td>1,250</td>
<td>Out-of-our health service</td>
<td>Validation</td>
<td>Out-of-hours Patient Questionnaire (OPQ)</td>
<td>Multidimensional (8)</td>
<td>1-Internal consistency 2-Structural validity 3-Reliability 4-Content validity</td>
<td>α = 0.70, inter-item and total item correlation (r &gt; 0.20)</td>
<td>1-Fair 2-Fair 3-Fair 4-Fair</td>
</tr>
<tr>
<td>Moret et al. 28 (2007)</td>
<td>2,684</td>
<td>Hospital</td>
<td>Adaptation</td>
<td>Echelle de Qualité des Soins en Hospitalisation (EQS-H scale)</td>
<td>Bidimensional</td>
<td>1-Internal consistency 2-Structural validity 3-Content validity 4-Measurement error</td>
<td>α &gt; 0.80, correlation inter-item (r &gt; 0.40) and item-total (r &lt; 0.40); RMSEA = 0.063</td>
<td>1-Good 2-Good 3-Fair 4-Good</td>
</tr>
<tr>
<td>Granado de la Orden et al. 44 (2007)</td>
<td>651</td>
<td>University hospital</td>
<td>Validation</td>
<td>Satisfacción de los Usuarios de las Consultas Externas (SUCE)</td>
<td>Bidimensional</td>
<td>1-Internal consistency 2-Structural validity 3-Hypotheses testing 4-Content validity</td>
<td>α &gt; 0.90; correlation with other instrument (r = 0.85); item-total correlation (r = 0.52-0.73)</td>
<td>1-Good 2-Good 3-Fair 4-Excellent</td>
</tr>
<tr>
<td>Shea et al. 37 (2008)</td>
<td>3,215</td>
<td>Private clinic</td>
<td>Adaptation</td>
<td>The Consumer Assessment of Healthcare Providers and Systems Surveys (CAHPS)</td>
<td>Multidimensional (7)</td>
<td>1-Internal consistency 2-Hypotheses testing 3-Content validity</td>
<td>α = 0.39-0.90</td>
<td>1-Fair 2-Poor 3-Fair</td>
</tr>
</tbody>
</table>

(continues)
Table 1 (continued)

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>N</th>
<th>Health care setting</th>
<th>Study design</th>
<th>Instrument used</th>
<th>Construct dimensions (*)</th>
<th>Measurement proprieties **</th>
<th>Psychometric values</th>
<th>Psychometric values</th>
<th>COSMIN scores **</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kristensson &amp; Ekwall 27</td>
<td>128</td>
<td>Hospital</td>
<td>Adaptation</td>
<td>Consumer Emergency Care Satisfaction Scale (CECSS)</td>
<td>Multidimensional (3)</td>
<td>1-Structural validity</td>
<td>$a &gt; 0.85$</td>
<td>1-Fair</td>
<td>1-Fair</td>
</tr>
<tr>
<td>(2008)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2-Internal consistency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clark et al. 51</td>
<td>1,415</td>
<td>Clinic and hospital (mental health services)</td>
<td>Validation</td>
<td>Consumer Perceptions of Care (CPC)</td>
<td>Multidimensional (4)</td>
<td>1-Internal consistency</td>
<td>$a &gt; 0.70$; correlations between instruments and different moments ($r = 0.34-0.56$)</td>
<td>1-Good</td>
<td>1-Good</td>
</tr>
<tr>
<td>(2008)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2-Structural validity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coluccia et al. 39</td>
<td>1,804</td>
<td>Hospital</td>
<td>Adaptation</td>
<td>Questionnaire of Quality Perception (QQP)</td>
<td>Multidimensional (4)</td>
<td>1-Structural validity</td>
<td>Correlations between factors 0.56-0.87; GFI = 0.92; RMSEA = 0.18; CFI = 0.96; NFI = 0.96</td>
<td>1-Excellent</td>
<td></td>
</tr>
<tr>
<td>(2009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arraras et al. 20</td>
<td>80</td>
<td>Hospital</td>
<td>Adaptation</td>
<td>European Organization for Research and Treatment of Cancer In-patient Questionnaire (EORTC IN-PATSAT 32)</td>
<td>Multidimensional (13)</td>
<td>1-Internal consistency</td>
<td>$a &gt; 0.70$; correlations with other instruments &gt; 0.40</td>
<td>1-Fair</td>
<td>1-Fair</td>
</tr>
<tr>
<td>(2009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2-Hypotheses testing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orden et al. 24</td>
<td>296</td>
<td>Hospital emergency</td>
<td>Validation</td>
<td>Satisfacción del Usuario con el Área de Urgencia Hospitalaria (SUJAUH)</td>
<td>Bidimensional</td>
<td>1-Internal consistency</td>
<td>$a &gt; 0.83$; correlation with other instrument and global satisfaction measures ($r = 0.37-0.51$); item-total correlation ($r = 0.56-0.71$)</td>
<td>1-Fair</td>
<td>1-Fair</td>
</tr>
<tr>
<td>(2010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2-Structural validity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Danielsen et al. 38</td>
<td>225</td>
<td>Clinic</td>
<td>Validation</td>
<td>Users’ experiences and satisfaction with casualty clinics</td>
<td>Multidimensional (4)</td>
<td>1-Internal consistency</td>
<td>$a &gt; 0.82$</td>
<td>1-Fair</td>
<td></td>
</tr>
<tr>
<td>(2010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continues)
Table 1 (continued)

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>N</th>
<th>Health care setting</th>
<th>Study design</th>
<th>Instrument used</th>
<th>Construct dimensions (*)</th>
<th>Measurement proprieties **</th>
<th>Psychometric values</th>
<th>COSMIN scores **</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ekwall &amp; Davis 2010</td>
<td>157</td>
<td>Hospital</td>
<td>Transcultural adaptation</td>
<td>Consumer Emergency Care Satisfaction Scale (CECSS)</td>
<td>Multidimensional (3)</td>
<td>1-Structural validity; 2-Internal consistency; 3-Cross-cultural validity</td>
<td>α &gt; 0.77</td>
<td>1-Fair</td>
</tr>
<tr>
<td>Charalamious 2010</td>
<td>194</td>
<td>Hospital (oncology center)</td>
<td>Transcultural adaptation</td>
<td>Risser Patient Satisfaction Scale (RPSS)</td>
<td>Multidimensional (3)</td>
<td>1-Structural validity; 2-Internal consistency; 3-Reliability; 4-Cross-cultural validity</td>
<td>α = 0.89; Kappa = 0.84; inter-item (r = 0.45) and item-total correlations (r = 0.51); sub-scale correlation (r = 0.55-0.70)</td>
<td>1-Fair</td>
</tr>
<tr>
<td>Erci &amp; Ciftcioglu 2010</td>
<td>381</td>
<td>Clinic (primary care)</td>
<td>Transcultural adaptation</td>
<td>Primary Care Satisfaction Survey for Women (PCSSW)</td>
<td>Multidimensional (3)</td>
<td>1-Internal consistency; 2-Structural validity; 3-Content validity; 4-Cross-cultural validity; 5-Reliability</td>
<td>α &gt; 0.80; item-total correlations (r = 0.43-0.73); test-retest correlation (r = 0.85)</td>
<td>1-Fair</td>
</tr>
<tr>
<td>Soufi et al. 2010</td>
<td>214</td>
<td>University hospital</td>
<td>Adaptation</td>
<td>Echelle de Qualité des Soins en Hospitalisation (EQS-H scale)</td>
<td>Bidimensional</td>
<td>1-Internal consistency; 2-Structural validity; 3-Content validity; 4-Cross-cultural validity</td>
<td>α &gt; 0.88; GFI = 0.96; RMSEA = 0.090; CFI = 0.93</td>
<td>1-Fair</td>
</tr>
<tr>
<td>Schröder et al. 2011</td>
<td>1,340</td>
<td>Hospital</td>
<td>Validation</td>
<td>Quality in Psychiatric Care – Outpatient (QPC-OIP)</td>
<td>Multidimensional (7)</td>
<td>1-Internal consistency; 2-Structural validity</td>
<td>α = 0.66; CFI = 1.0; RMSEA = 0.036; SRMR = 0.053; inter-scales correlations (r = 0.35-0.67)</td>
<td>1-Excellent 2-Excellent</td>
</tr>
<tr>
<td>Webster et al. 2011</td>
<td>486</td>
<td>Hospital</td>
<td>Validation</td>
<td>Patient Assessment of Healthcare for Inpatient Care (I-PAHC); Patient Assessment of Healthcare for Outpatient Care (O-PAHC)</td>
<td>Multidimensional (5)</td>
<td>1-Internal consistency; 2-Structural validity; 3-Content validity</td>
<td>α &gt; 0.70; inter-scales and item correlations (r = 0.2-0.6)</td>
<td>1-Good 2-Good 3-Poor</td>
</tr>
</tbody>
</table>

(continues)
<table>
<thead>
<tr>
<th>Study (year)</th>
<th>N</th>
<th>Health care setting</th>
<th>Study design</th>
<th>Instrument used</th>
<th>Construct dimensions (*)</th>
<th>Measurement properties **</th>
<th>Psychometric values</th>
<th>COSMIN scores **</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latour et al. 52 (2012)</td>
<td>441</td>
<td>University hospital (neonatal intensive care)</td>
<td>Validation</td>
<td>Empowerment of Parents in the Intensive Care-Neonatology (EMPATHIC-N)</td>
<td>Multidimensional (5)</td>
<td>1-Internal consistency 2-Structural validity 3-Reliability 4-Hypotheses testing 5-Content validity</td>
<td>α &gt; 0.82; correlation with other instrument and global satisfaction measures (r = 0.37-0.51)</td>
<td>1-Good 2-Good 3-Good 4-Fair 5-Good</td>
</tr>
<tr>
<td>Evans et al. 23 (2012)</td>
<td>397</td>
<td>Clinic (mental health)</td>
<td>Validation</td>
<td>Views on Inpatient Care (VOICE)</td>
<td>Multidimensional (7)</td>
<td>1-Structural validity 2-Internal consistency 3-Content validity 4-Hypotheses testing 5-Reliability validity</td>
<td>α &gt; 0.92; correlation with other instrument (r = 0.82); Kappa (κ = 0.88)</td>
<td>1-Poor 2-Poor 3-Fair 4-Fair 5-Fair</td>
</tr>
<tr>
<td>Manaf et al. 43 (2012)</td>
<td>424</td>
<td>Clinic</td>
<td>Validation</td>
<td>Outpatient satisfaction instrument</td>
<td>Bidimensional</td>
<td>1-Internal consistency 2-Structural validity 3-Hypotheses testing 4-Content validity</td>
<td>α &gt; 0.94; correlation with other instruments and overall satisfaction measure (r = 0.59-0.62)</td>
<td>1-Excellent 2-Excellent 3-Excellent 4-Excellent</td>
</tr>
<tr>
<td>Milutinović et al. 32 (2012)</td>
<td>235</td>
<td>Hospital</td>
<td>Transcultural adaptation</td>
<td>Patient Satisfaction Nursing Care Quality Questionnaire</td>
<td>Unidimensional</td>
<td>1-Structural validity 2-Internal consistency 3-Cross-cultural validity</td>
<td>α &gt; 0.90; inter-item correlation and item total correlation (r =0.52-0.92); GFI (AUC = 0.82)</td>
<td>1-Good 2-Good 3-Good</td>
</tr>
<tr>
<td>Ottonello et al. 21 (2012)</td>
<td>7,093</td>
<td>Hospital (emergency care)</td>
<td>Adaptation</td>
<td>SAT-16</td>
<td>Bidimensional</td>
<td>1-Structural validity 2-Internal consistency</td>
<td>α &gt; 0.99; correlation between item (r &lt; 0.30); Rasch reliability values (Item separation index: 18.1 and 20.7; item separation reliability: 1; person separation index: 0.70)</td>
<td>1-Excellent 2-Excellent</td>
</tr>
</tbody>
</table>

(continues)
Table 1 (continued)

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>N</th>
<th>Health care setting</th>
<th>Study design</th>
<th>Instrument used</th>
<th>Construct dimensions (*)</th>
<th>Measurement proprieties **</th>
<th>Psychometric values</th>
<th>COSMIN scores **</th>
</tr>
</thead>
<tbody>
<tr>
<td>Castello et al. 54 (2012)</td>
<td>483</td>
<td>Hospital</td>
<td>Adaptation</td>
<td>Satisfaction Scale with Primary Health Care Services</td>
<td>Multidimensional (4)</td>
<td>1-Structural validity</td>
<td>$\alpha &gt; 0.75$; RMSEA = 0.09; GFI = 0.92; AGFI = 0.87; NFI = 0.94; CFI = 0.95</td>
<td>1-Fair</td>
</tr>
<tr>
<td>Dyer et al. 55 (2012)</td>
<td>21,318</td>
<td>Hospital/Clinic</td>
<td>Adaptation</td>
<td>Consumer Assessment of Healthcare Providers and Systems (CAHPS-CG)</td>
<td>Multidimensional (3)</td>
<td>1-Structural validity 2-Internal consistency 3-Reliability</td>
<td>$\alpha &gt; 0.75$; RMSEA = 0.05; SRMR = 0.04; CFI = 0.97; ICC = 0.01-0.11</td>
<td>1-Good</td>
</tr>
<tr>
<td>Vanti et al. 35 (2013)</td>
<td>317</td>
<td>Hospital</td>
<td>Transcultural adaptation</td>
<td>Physical Therapy Patient Satisfaction Questionnaire (PTPSQ-I)</td>
<td>Multidimensional (3)</td>
<td>1-Structural validity 2-Internal consistency 3-Cross-cultural validity 4-Hypotheses testing</td>
<td>$\alpha &gt; 0.90$; correlations between item and global perceive measures ($r = 0.1-0.42$)</td>
<td>1-Fair</td>
</tr>
<tr>
<td>Vanti et al. 56 (2013)</td>
<td>354</td>
<td>Hospital</td>
<td>Transcultural adaptation</td>
<td>Physical Therapy Patient Satisfaction Questionnaire (PTPSQ-I)</td>
<td>Multidimensional (4)</td>
<td>1-Structural validity 2-Internal consistency 3-Cross-cultural validity 4-Reliability 5-Hypotheses testing</td>
<td>$\alpha &gt; 0.75$; ICC = 0.62-0.93; correlations between item and global perceive measures ($r = 0.28-0.42$)</td>
<td>1-Fair</td>
</tr>
</tbody>
</table>

* Number of dimensions in the instrument (construct representation); ** The methodological classification scores for each item of the Measurement Properties column are demonstrated by the respective number in the COSMIN column.

$\alpha$: Cronbach’s alpha; AGFI: adjusted goodness of index; AUC: area under the ROC curve; CFI: comparative fit index; GFI: goodness of fit index; ICC: intra-class correlation; NFI: normed fit index; RMSEA: root mean square error of approximation; SRMR: standardized root mean residual.

Satisfaction construct

One can argue that there is almost a consensus between studies on the multidimensional characteristic of the satisfaction construct (Table 1). Only three studies showed a unidimensional structure for the tested instrument 22,32,48. However, there is not a clear pattern of what the essential dimensions (factors) that must be included in a satisfaction measurement are. The number of dimension/factors in the instruments scanned ranged from 1 to 13.

It was also possible to observe a trend showing that more than one dimension should be used to develop or adapt an instrument (Table 1). The main dimensions were: patient–healthcare professional interactions, physical environment and internal management process 22,23,32,48. Only services with some different characteristics (home care, out-of-hours, psychiatric and pediatric services, and more multidisciplinary complex services) usually utilized other specific dimensions 22,23,49.
Discussion

Considering the focus on patient-centered care has been proposed all around the world, the development or adaption of satisfaction surveys has been the most common method to evaluate patient satisfaction. However, the very high specificity of the different types of care makes the development of a single measurement scale transposed into all services more complex. The findings of this revision agree with previous studies showing that most surveys demonstrated “user satisfaction in health care services” as a multifactorial construct, but until now there is not a reference or gold standard instrument for the measurement of such a concept.

Previously, Boyer et al. showed, in their systematic review about patient satisfaction surveys in psychiatric services, a wide variation in the number of dimensions of the instruments scanned. Even though our review had screened instruments applied to different healthcare settings, the results support these findings, and the number of dimensions in the instruments scanned ranged from 1 to 13.

A good theory-based approach is the first step to define any construct. When it comes to user satisfaction, it has already been shown that the essential aspects are: patient-professional interactions, physical environment and internal management processes. However, services with some different characteristics (home care, out-of-hours, psychiatric and pediatric services, and more multidisciplinary complex services) usually utilized other specific elements, such as phone contact or clear information to the caregiver.

The complexity of this construct can make it difficult to develop measurement instruments and to define their real contribution for healthcare quality and management. It must be understood that the concept of satisfaction can change across many cultures, and different variables may be related depending on patient profile or service. Almost all clinicians and patients agree that clinical skill, rapport and health-related communication behavior are key elements for quality and satisfaction. However, our findings demonstrated that all multidimensional instruments utilized sub-scales that were linked with these key elements.

The methodological quality of the questionnaires and their validation process are important topics that should be highlighted. All instruments found were developed in the English language and there were only six cross-cultural adaptation papers. The adaptation process to other languages and cultures is not simple and the scrutiny made by different researches on distinct sociocultural contexts would help to develop a reliable instrument.

The psychometric properties of the questionnaires are just one point of the validation or adaptation process and other steps should be taken. Our results showed that structural validity and internal consistency were the main measurement properties described by the studies, but none of them had already undergone all other steps proposed by the COSMIN checklist for validation: reliability, measurement error, content validity, hypotheses testing, criterion validity, cross-cultural validity and responsiveness.

Discussion about what drives users’ satisfaction is necessary to guide the development of new surveys and classic confirmatory analysis. Assessing this issue is quite important to choose an instrument that has already been scrutinized by several authors and crossed many important validation steps. However, this approach has still proved to be insufficient in the field of quality of care, especially on patient satisfaction questionnaires.

Unfortunately, until now there are few guidelines to evaluate instruments’ validation processes and help to define the most important steps for the development of a new instrument, which led to the large quantity of instruments found. This study used the COSMIN guide to check the quality of each measurement property described on instruments validation process and psychometric properties. This analysis demonstrated that most of the instrument achieved a medium score (”good” or “fair”). The comparison with other studies is difficult because no other systematic review with similar characteristics was found in the literature.

Even though patient satisfaction has been widely studied in healthcare, no gold standard instrument or even a validated instrument was found for many different contexts. The option to develop new instruments by the healthcare organizations can reflect the specificity of each culture or health settings, but, at the same time, this phenomenon can show a gap in state-of-the-art approaches in this area.

It should be highlighted that users’ satisfaction is just one facet of quality of care concept and the dimension related to this topic is the patient-centered care. Satisfaction studies can help some local health policies because they can show a tendency of presence of quality on care. It had already been shown that improvement in actions to understand patient preferences or...
patient-physician communication can impact outcomes 64.

Rathert et al. 65 reported in their systematic review that patient centered care actions may affect patients over time in ambulatory settings. Moreover, to understand how these concepts are related to patient health outcomes can help to determine the effectiveness of patient centered care.

This review presents limitations and even through some search equations had been tested, the final search strategy assumed restricted headings and only papers published since 2002. The search strategy used the headings merging the terms “patient centered care” and “patient satisfaction”. In addition, the term ‘psychometric properties’ was chosen to capture all kinds of possible analyses. This choice could lead to excluding articles classified as validity or reliability, although previous tests have shown that it were included. The temporal limitation was included because the literature on patient centered care as a dimension of quality in health care has become more widespread recently 3,4. Hence, the development of instruments for patient satisfaction based has also proved to be more recent. Another limitation of the present study is caused by access problems to bibliographic databases, whereby important databases such as EMBASE were not screened. Furthermore, the search strategy faced some difficulties to find studies from different areas of healthcare, such as physiotherapy, odontology or nutrition. This can give rise to some difficulties when it comes to the interpretation of the satisfaction dimensions for all health fields and the structure of the main instruments.

Conclusion

This review showed that there is not a consensus about a reference instrument (gold standard) for patient satisfaction measurement in healthcare. However, the main dimensions of many instruments utilized are patient-professional interactions, physical environment and internal management processes. The measurement instruments analyzed have shown a medium score of methodological quality by the COSMIN checklist. This construct presents many options of measurement, but the utilization of an instrument to measure patient satisfaction of health services should consider the quality of the validation process of each questionnaire. Finally, it seems clear that the “patient satisfaction construct” should be measured considering a multidimensional approach, because even the patient’s clinical profile, such as health service environment or professional behavior, can be determinants for patient satisfaction. Furthermore, within the context of shared decision-making and preference sensitive care, it is important to highlight the need to better understand the role of patient experience and satisfaction as a measure of quality of care 12,66.
Resumen

Los cuestionarios para la satisfacción del paciente pueden ser una buena opción de mejorar la calidad y discutir la atención centrada en el paciente. El objetivo era llevar a cabo una revisión sistemática, en relación con los instrumentos previamente validados para medir la satisfacción de los pacientes. La revisión sistemática realizó búsquedas en las bases de datos MEDLINE/PubMed, LILACS, SciELO, Scopus y Web of Knowledge. La estrategia de búsqueda fue: “Patient Satisfaction” AND “Patient centered care” AND “Healthcare survey OR Satisfaction questionnaire” AND “Psychometric properties”. Se incluyeron treinta y siete estudios y casi todos han demostrado que la satisfacción es multidimensional. Se encontraron 34 instrumentos y la mayoría contaba con las dimensiones: interacción paciente-profesional de la salud, entorno físico y procesos administrativos. La puntuación COSMIN en calidad metodológica indica que la mayoría de los estudios tiene un valor bueno o medio. No existe un estándar clave para la medición de la satisfacción del paciente, pero algunas dimensiones son esenciales para este constructo.

Calidad de la Atención de Salud; Atención Dirigida al Paciente; Satisfacción del Paciente; Psicometría

Contributors

R. S. Almeida and M. Martins were responsible for the design and search strategy. R. S. Almeida was responsible for conducting the search and provided input into the data analysis. Data interpretation and the writing of the manuscript were carried out by R. S. Almeida, M. Martins and S. Bourliataux-Lajoinie.

Acknowledgments

The authors would like to thank the National School of Public Health, Oswaldo Cruz Foundation (ENSP/Fiocruz) for their structural and financial support in carrying out the research and for help in finding all papers.

References


Submitted on 19/Feb/2014
Final version resubmitted on 14/Oct/2014
Approved on 20/Oct/2014