emeraldinsight

International Journal of Health Care Quality Assurance

Patient-perceived hospital service quality: an empirical assessment

YOGESH P PAI, Satyanarayana T Chary, Rashmi Yogesh Pai,

Article information:

To cite this document:
YOGESH P PAI, Satyanarayana T Chary, Rashmi Yogesh Pai, "Patient-perceived hospital service quality: an empirical assessment", International Journal of Health Care Quality Assurance, https://doi.org/10.1108/IJHCQA-04-2017-0064
Permanent link to this document:
https://doi.org/10.1108/IJHCQA-04-2017-0064

Downloaded on: 24 January 2018, At: 20:59 (PT)

References: this document contains references to 0 other documents.

To copy this document: permissions@emeraldinsight.com

Access to this document was granted through an Emerald subscription provided by emerald-srm:178665 []

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Patient-perceived hospital service quality: an empirical assessment

Author 1: Yogesh Pai P, B.E., M.B.A., M.Phil., Ph.D., Associate Professor, School of Management, Manipal, India, Telephone: +91-9880530633, e-mail: yogesh.pai@manipal.edu, yogeshpai.p@gmail.com

Author 2: Satyanarayana T Chary, M. Com, M.F.M., M.B.A., Ph.D., Professor, Department of Business Management, Telangana University, Nizamabad, India, Telephone: +91-9704599931, e-mail: tsnchary@gmail.com

Author 3: Rashmi Yogesh Pai, B.Sc., M.B.A., Ph.D., Assistant Professor, Poornaprajna Institute of Management, Udupi, India, Telephone: +91-9880530632, e-mail: rashyogipai@gmail.com

Corresponding author: Yogesh Pai P
Corresponding Author's E-mail: yogesh.pai@manipal.edu, yogeshpai.p@gmail.com

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Structured Abstract:
Purpose: Our purpose was to appraise Pai and Chary's (2016) conceptual framework for measuring patient-perceived hospital service quality.
Design/Methodology/Approach: A structured questionnaire was used to obtain data from teaching, public and corporate hospital patients. Several tests were conducted to assess the instrument's reliability and validity. Pai and Chary's (2016) nine dimensions for measuring hospital service quality (HSQ) were examined.
Findings: Tests confirm that Pai and Chary's (2016) conceptual framework is reliable and valid. Our study also establishes that nine dimensions measure hospital service quality (HSQ).
Practical implications: The framework empowers managers to assess service quality in any hospital settings: corporate, public and teaching, using an approach that is superior to existing hospital service quality scales.
Originality/Value: This article helps researchers and practitioners to assess hospital service quality from patient perspectives in any hospital setting.
Keywords: Reliability; Scale assessment; Hospital service quality; Validity; India.

Article Classification: Research

Received – 8th Sep 2016
Revised – 15th April 2017
Accepted – 1st June 2017

Introduction

Healthcare is considered a 'credence' good – an offering that consumers will never be able to evaluate owing to deficient medical knowledge (Bloom and Reeve, 1990). Additionally, conceptualising and measuring service quality in a healthcare setting is more important and simultaneously more complex (Taner and Antony, 2006). However, researchers need to come up with ways to measure healthcare service quality, because unless we measure we cannot manage and improve healthcare services (Lohr, 2015). The literature indicates that there is a variability and confusion in how quality is conceptualized and operationalized (Sower *et al.,* 2001). Many researchers attempt to define and conceptualize hospital service quality. Pai and Chary (2013),

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

for example, show that the Parasuraman *et al.,* (1985, 1988) SERVQUAL/modified SERVQUAL questionnaire is conventionally practiced in healthcare. They also pointed out studies where SERVQUAL items have not loaded onto respective dimensions highlighting a five-component structure that is lacking, suggesting a new questionnaire needs developing. New instruments, designed for healthcare setting: PRIVHEALTHQUAL (Ramsaran-Fowdar, 2008) for private and PubHosQual (Aagja and Garg, 2010) for public settings, have emerged. Studies adopting these instruments are scarce because these scales are hospital specific and not a general scale that measures hospital service quality in any hospital context. An instrument that could measure hospital service quality (HSQ) in any hospital settings has gained importance. Pai and Chary (2016) proposed their conceptual framework for measuring hospital service quality using nine dimensions. Their framework, unlike other instruments such as SERVQUAL (Parasuraman *et al.,* 1988), PRIVHEALTHQUAL (Ramsaran-Fowdar, 2008) and PubHosQual (Aagja and Garg, 2010) has an additional dimension (relationship) in line with researchers such as Carman (1990) and Reynoso and Moore (1995), who suggested adding service specific dimensions to SERVQUAL, there exists no empirical studies to support or refute their conceptual framework. As there are no research studies that used Pai and Chary's framework (2016), our purpose is to empirically appraise this framework and addresses literature gaps.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Objectives

We conducted our study with the objective to test the Pai and Chary (2016) conceptual framework's validity and reliability. We also aim to assess HSQ through nine dimensions: (i) healthscape; (ii) personnel; (iii) hospital image; (iv) trustworthiness; (v) clinical process; (vi) communication; (vii) relationship; (viii) personalization and (ix) administrative procedures.

Structured questionnaire

The questionnaire included:

1. Respondent demographics – gender, age, occupation, educational level, marital status and income.
2. Sixty-six items covering nine dimensions (Pai and Chary, 2016).
3. Four items measuring respondents' hospital service quality perception.
4. Twelve items that measured respondent satisfaction and behaviour.

A Likert (1932) scale, having a balanced rating with equal categories above and below the midpoint, was used with each item; anchored with verbal statements like 'strongly disagree', 'disagree', 'neither agree or disagree', 'agree', 'strongly agree'. All items were phrased positively (Easterby-Smith *et al.,* 2002; Pai and Chary, 2013; Parasuraman *et al.,* 1994). Sixty-six hospital service quality items were categorized into nine dimensions: (i) healthscape (15 items); (ii) personnel (11); (iii) hospital image (5); (iv) trustworthiness (5); (v) clinical process (6); (vi) communication (9); (vii) relationship (3); (viii) personalization (3); and (ix) administrative procedures (9 items). Four perceived service quality, twelve satisfaction and behavioural intention items were included (Appendix 1).

Piloting

Piloting a new instrument is imperative (DeVellis, 1991; Kirchhoff, 1999). Before the pilot test is performed, it is advisable to obtain the original evaluation (DeVellis, 1991). Andres (2012) recommends that conducting a pilot study by involving colleagues, friends and family members

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

who assume an audience role. Our questionnaire was subjected to the pre-testing process by circulating it among ten colleagues, friends and family members. In any research, respondents often misunderstand words or concepts. Although communication difficulties exist, respondents still provide legitimate answers to survey questions (Clark and Schober, 1992; Tourangeau *et al.,* 2000). To overcome these problems, Collins (2003, p.230) suggested cognitively testing survey questions that helps 'to identify how and where the question fails to achieve its measurement purpose'. Although there are various cognitive methods that have been developed and applied to instrument testing, such as cognitive interviewing, paraphrasing, card sorts, vignettes, confidence ratings and response latency timing (Czaja, 1998; Forsyth and Lessler, 1991; Jobe and Mingay, 1991); cognitive interviewing is becoming widespread (Schwarz, 1997). Cognitive interviewing (involving two methods: think aloud interviewing and probing) and paraphrasing were used in testing. On completion, the instrument was tried in a hospital with patients as respondents. According to Lackey and Wingate (1998), pilot testing a newly developed instrument should be undertaken with respondents selected from the same population from which the subjects in the major study will be selected. Consequently, the questionnaire was tested with respondents from teaching, corporate and public hospitals. Pre-testing ensured correct phrasing, format, length and question sequence. Pretesting was done with 30 respondents using Hertzog's (2008) guidance. The questionnaire was corrected after feedback.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Data collection

Data were collected in three hospitals – teaching, corporate and public in Karnataka. India has 29 states and seven union territories (Ashok, 2014), and every state has a specific language. Although India is an English-speaking country, Kannada is Karnataka's local language. Hence, the questionnaires were in Kannada and translated into the Malayalam language as few hospitals had Malayalam speaking patients in significant numbers owing to hospital's proximity to another state (Kerala), where the official language is Malayalam. Consequently, questionnaires were administered in three languages: English; Kannada and Malayalam. There are several studies in which the questionnaire was administered in different languages specific to the country studied; e.g., English and Maltese (Camilleri and O'Callaghan, 1998); English and Arabic (Jabnoun and Chakar, 2003); English and Turkish (Kara *et al.,* 2005); English and Gujrati (Aagja and Garg, 2010); and English and Bengali (Akter *et al.,* 2008). The Kannada/Malayalam versions were created through careful translation and back-translation techniques (Candell and Hulin, 1987; McGorr, 2000). First, the questionnaire was translated into Kannada/Malayalam and then the Kannada/Malayalam items were back-translated into English by a bilingual expert to ensure that the original content was kept intact. In translating the scale items into Kannada/Malayalam, Malinowski's (1935) four-step translation technique was implemented:

1. An interlinear, or word-by-word, translation.
2. A 'free' translation in which clarifying terms, conjunctions, etc., are added and words reinterpreted.
3. Analysing and collating two translations leading to:
4. A contextual specification of meaning.

An attempt was made to remove discrepancies between English and Kannada/Malayalam questionnaire versions. No individual items were found to be problematic in translation. Previous studies reported backward translation; e.g., English to Bengali (Akter *et al.,* 2008; Andaleeb,

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

2000); English to Japanese (Amira, 2008); English to Arabic (Mostafa, 2005); English to Hindi (Rao *et al.,* 2006). Accordingly, our instrument was translated from English to Kannada and Malayalam languages.

Sample

The study population was defined as all patients 18 years or older with a stable mental and clinical condition during the data collection procedure. A stratified random sampling was adopted for the study. As the name implies, there is a stratification or segregation process followed by randomly selecting subjects from each stratum. The population is first divided into mutually exclusive groups which are relevant, appropriate and meaningful in the study context. The teaching, corporate and public hospitals constitute three strata. Hospitals were selected using systematic sampling and proportionate stratified sampling. The sampling frame was prepared for three strata and ten per cent from each stratum: three hospitals each from teaching and public hospitals and four corporate hospitals, totalling ten hospitals. Respondent from ten hospitals were chosen randomly (every fifth element in the population), both inpatients and outpatients yielding 602 respondents. The respondent's demographic profile is depicted in the Table I, comprising males (44%). Their sample's age ranged from 18 to 67+ years.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Table I here

Reliability and validity

Instruments designed to measure a specific concept should measure what is set out to measure. Reliability and validity are the two main criteria for measuring how good measures are in any research instrument. Reliability indicates stability and consistency with which the instrument measures the concept, thus assessing a measure's goodness (Sekaran and Bougie, 2010). It is important to calculate scale reliability, which refers to the extent to which a scale can reproduce the same results in repeated trials (Hair *et al.,* 2003). Consistency can be examined through the inter-item consistency reliability and split-half-reliability tests (Sekaran and Bougie, 2010). The internal consistency reliability test is acceptable when the reliability coefficient exceeds Nunnally's (1978) 0.7 reliability criterion.

The questionnaire included 82 items (Appendix 1). For internal consistency, the most popular test is coefficient alpha (Cronbach). For Pai and Chary's (2016) 66 items, the value is 0.965 and 0.972 for the 82-item instrument. The higher the coefficients, the better the measuring instrument (Sekaran and Bougie, 2010). Another technique for testing internal consistency is the split-half technique, where items constituting the instrument are divided into two halves and the resulting half scores are correlated (Malhotra and Dash, 2012). For the 66-item version, Cronbach's Alpha was 0.93 and 0.947, correlation between forms was 0.793, Spearman–Brown Coefficient for equal length and unequal length is 0.884, Guttman Split-Half Coefficient is 0.878. For the 82-item questionnaire, Cronbach's Alpha was 0.944 and 0.956, correlation between forms was 0.821, Spearman–Brown Coefficient for equal length and unequal length is 0.902, Guttman Split-Half Coefficient is 0.897. We can conclude, therefore, that the instrument has adequate reliability. Cronbach's alpha is a widely-used reliability coefficient that assesses the entire scale's internal consistency and averages all possible split-half coefficients resulting from different ways of splitting the scale items ranging from 0 to 1. According to Malhotra and Dash (2012), it is appropriate to use internal consistency reliability for each dimension, if several items are used to measure each dimension. In our study, patient-perceived hospital service quality is

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

measured using different dimensions, each is measured by several items, hence we used Cronbach's alpha to measure internal consistency. Cronbach's alpha for the twelve constructs considered in the study was measured and values for each construct are 0.70 and above, indicating a strong reliability (Table II).

Table II here

Although reliability is a necessary contributor to validity, it is not a sufficient condition for validity (Sekaran and Bougie, 2010). Therefore, validity becomes equally important. Validity is 'the extent to which a rating scale truly reflects the underlying variable that it attempts to measure' (Parasuraman *et al.,* 2004, p. 294). Content validity, criterion-related validity and construct validity are generally discussed in research articles.

Content validity

Content validity is the degree to which items adequately represents all relevant items under study (Cooper *et al.,* 2012), determined using judgment and panel evaluation. For an attitudinal scale, content validity is an overall criterion that can be assessed only though a researcher's subjective judgment (Parasuraman *et al.,* 2004). We exercised judgment through carefully defining and analysing conceptual and empirical frameworks through an extensive literature review. Judges can attest to content validity (Sekaran and Bougie, 2010), so the questionnaire was subjected to expert review by practitioners and academics. Content/face validity was confirmed; i.e., the proposed measurement instrument measures the intended construct. The items for the current study were chosen from the literature (Sekaran and Bougie, 2010). Scales were refined using a pilot study. Patients revealed they had no difficulty understanding the questionnaire items indicating and confirming face validity (Arasli *et al.,* 2008). All steps ensured that the instrument possesses face validity. Although implementing a scale based on only face validity claims has a far greater impact than item's random deletion through mere formal scale refinement methods (Finn and Kayande, 2004), nevertheless, other validity tests are also conducted.

Criterion related validity (Ping, 2004) emphasizes that, for new measures, criterion validity should be assessed, which is established when the measure differentiates individuals on a criterion it is expected to predict. Criterion related validity can be established by concurrent or predictive validity (Sekaran and Bougie, 2010). Predictive validity indicates the instrument's ability to differentiate among individuals with reference to a future criterion (Sekaran and Bougie, 2010). The researcher needs to ensure that the validity criterion used is valid and the intended measure is judged on four qualities: relevance; freedom from bias; reliability; and availability (Cooper *et al.,* 2012). Criterion-related validity was established using predictive validity, adopting Jabnoun and Chakar's (2003) technique, who correlated their service quality dimensions with overall service quality. Individual item scores on nine dimensions (listed earlier) were summed to obtain overall scores for each respondent. Scores were then correlated with a summated perceived service quality, satisfaction and behavioural intentions scale. All items were captured and measured using Likert scales with a five-point response format. A higher score indicated a more favourable response (Table III). All correlations were positive and statistically significant at 0.001 level, which establishes predictive validity.

Table III here

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Concurrent validity was assessed by examining the association between perceived service quality and a related variable, patient satisfaction (Hansen *et al.,* 2008). Patient satisfaction is a distant but related construct to perceived service quality (Choi *et al.,* 2004; Rao *et al.,* 2006). Although there are differences between patient satisfaction and perceived service quality, they are related and a moderate correlation is expected (Al-Qatari and Haran, 1999). Crompton and Love (1995) stressed that the two constructs are likely to be positively correlated, but unlikely to be linear. The correlation between perceived service quality and patient satisfaction is 0.633 (Table III), which indicates concurrent validity. Concurrent validity assessment is conducted in line with Duggirala *et al.,* (2008) such that in the present context, perceived service quality (PSQ) is chosen as the nine-dimension's outcome and a bivariate correlation analysis was carried out among all constructs, which have significant positive correlations with the criterion chosen, namely, PSQ (range 0.414 to 0.587, $p < 0.001$). According to Sekaran and Bougie (2010), construct validity testifies how well the results obtained from a measure fit the theories around which the test is designed and assessed through convergent and discriminant validity. Convergent validity is the degree to which scores on one scale correlate with scores on other scales designed to assess the same construct (Cooper *et al.,* 2012). Convergent validity is also assessed by calculating one-way analysis of variance (ANOVA) (Aagja and Garg, 2010; Bahia and Nantel, 2000; Parasuraman *et al.,* 1988; Wong *et al.,* 2001). The association between healthscape, personnel, hospital image, trustworthiness, clinical care process, communication, relationship, personalization and administrative procedures and PSQ was significant at less than 1% (Table IV), which means that the group differences are significant; i.e., good convergent validity.

Table IV here

Alternate methods for establishing criterion and construct validity
Campbell and Fiske's (1959) procedure has been used as an alternate method for establishing criterion and construct validity, adopted in similar studies such as Boshoff and Terblanche, 1997; Kaul 2007. Convergent validity (Table V) is confirmed as there is a high correlation between scale items and PSQ (correlation coefficient, $r = 0.665$, $p<0.001$). To assess predictive validity, respondents were asked whether they intended to return to the same hospital (repurchase intentions). Results confirm predictive validity ($r = 0.621$, $p< 0.001$).

Table V here

Empirically validating nine patient-perceived hospital service quality dimensions
We propose the following hypothesis:

H_o: The heathscape, personnel, hospital image, trustworthiness, clinical process, communication, relationship, personalization and administrative procedure domains do not predict hospital service quality.

To test our hypothesis, a multiple regression approach is adopted with the nine dimensions as independent variables and PSQ as the dependent variable. Variables were examined for multicollinearity and autocorrelation before the regression analysis was run. Multicollinearity exists when there is a strong correlation between two or more predictors in a regression model

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

leading to a regression model that fits the data, but no variable had a significant impact in predicting the dependent variable. In SPSS, there are collinearity diagnostics that are variance inflation factor (VIF) and tolerance statistics. The VIF indicates whether the predictor has a strong linear relationship with other predictor(s), while the tolerance statistic reciprocates the variance inflation factor (VIF). We used the Durbin-Watson test to identify any serial correlations between errors (Table VI and VII).

Table VI here

The tolerance values need to be greater than 0.1 and the VIF values should be less than 10 (Ho, 2006). From the tolerance and VIF values shown in Table VI, multicollinearity does not exist in the predictor values. Hence, the nine independent variables influence on dependent variables is analysed. The regression model is formed using the statistical (stepwise) regression method taking all nine independent variables together and examining their combined effect on the dependent variable. In the statistical (stepwise) regression model, the predictor variables' order is solely based on statistical criteria and variables that correlate most strongly with the dependent variable will be given entry priority. Our regression models require no references to theoretical considerations, and are primarily used when the researcher is unsure about the independent variable's predictive power (Ho, 2006). This method is apt for the present study and therefore stepwise regression is adopted, which includes forward inclusion, backward elimination and stepwise solution (Ho, 2006; Malhotra and Dash, 2012). To identify an optimal regression equation, one needs to compute combinational solutions in which all possible combinations are examined (Malhotra and Dash, 2012). The optimal regression equation in our study is computed using all possible combinations. However, the stepwise regression method is preferable to the forward method, as they reduce Type II errors (i.e., missing a predictor that does in fact predict the outcome) (Field, 2013). Regression analysis results are shown in Table VII.

Table VII here

The F value tests how well the regression model fits the data (Ho, 2006). In case the probability associated with the F statistics is small, the hypothesis that R-square is equal to zero is rejected. In our case, F is 71.65, $p < 0.001$ (Table VII); i.e., the hypothesis that there is no linear relationship between the predictor and the dependent variables is rejected. The R statistic indicates that the correlation coefficient as 0.677 and the R-square value is 0.458. For our sample, the predictor variables explained 45.8 per cent of the variance in the dependent variable. Hence, the null hypothesis H_o is rejected. In Table VII, Durbin-Watson's d = 1.987. Durbin-Watson can vary between 0 and 4; i.e., 2 implies that the residuals are uncorrelated. If the value is greater than 2, then it indicates a negative correlation between adjacent residuals; whereas a value below 2 indicates a positive correlation (Field, 2013). Thus, Durbin-Watson statistic tests serial correlations between errors, a metric that assesses another linear model assumption. Field (2013) suggests that the stepwise method should be avoided except for exploratory model building. He recommends cross-validating the model if the stepwise method is used, suggesting two cross-validation methods: adjusted R^2 and data splitting.

Cross-validating the model using adjusted R square (R^2)

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

The adjusted R^2 shows how well our model generalizes. Although the adjusted R^2 should be the same or very close to R^2, it differs generally. This difference is known as shrinkage, which = R^2 - Adjusted R^2. As per model summary (Table VII) shrinkage is equal to 0.458 - 0.451 = 0.007 (about 0.7 per cent); i.e., the model's shrinkage is small. Although there are no shrinkage guidelines in the literature, it is valuable cross-validation such that one can be confident in the equation's generalizability if they are similar (Pedhazur, 1997). According to Field (2013), shrinkage means that if the model were derived from the population rather than a sample, then it would account for approximately 0.7 per cent less variance in the outcome. Stein's formula helps us to calculate adjusted R^2 and tells us how well the model cross-validates (Stevens, 2002), which is computed as: Adjusted R^2 = 1-[(n-1/n-k-1) (n-2/n-k-2) (n+1/n)] (1- R^2) where R^2 is the unadjusted value, n is total participants and k is total predictors in the model. Substituting R^2 as 0.458 (Table VII), n = 602 and k=9 in the above formula, we get 0.441, which is computed as 1-[(602-1/602-9-1) (602-2/602-9-2) (602+1/602)] (1- 0.458), which = 0.441. This value is like the observed 0.458, indicating that cross-validity is acceptable (Field, 2013, p.312).

Cross-validating the model using data splitting

The other cross-validation method is the data splitting approach. Cross-validation is a validity test that examines whether the regression model continues to hold on comparable data not used in the original estimation (Malhotra and Dash, 2012). In a cross-validation procedure, we estimate the regression model for the entire data, which are split into two parts: estimation and validation samples. The estimation sample incudes 50-90% of the total sample. The regression model is estimated using the estimation sample only. This model is compared with the entire sample. The estimated model is applied to the data in the validation sample to predict the dependent variable. In case there is a large discrepancy between R^2 for the smaller and larger samples, this indicate over fitting and weak generalizability (Tabachnick and Fidell, 2014). In other words, if there is a significant difference between R^2 for the larger sample (estimation sample) and smaller sample (validation sample) then there is weak generalizability. Using SPSS, the step-wise regression on a random selection is adopted as suggested by Field (2013). Cases are randomly divided into two parts for cross-validation. Creating a variable that contains information about the estimation sample and validation samples using 0 and 1 to indicate the estimation sample and validation sample. We split the data into two samples: estimation and validation samples with 325 (53.9%) and 277 (46.1%) respondents/samples using Malhotra and Dash, (2012) guidance that estimation sample includes 50-90% of the total sample. Stepwise regression is conducted on the split data; i.e., estimation and validation sample, and the R^2 values are compared with the original model (Malhotra and Dash, 2012; Tabachnick and Fidell, 2014). The overall relationship between the dependent variable (PSQ) and the independent variable (nine dimensions) must be statistically significant for both validation analyses. Table VIII and IX indicate that significance is achieved. The R^2 for each validation must be within ±5% (plus or minus 5%) of the model's R^2 using the full sample. The entire sample's R^2 is 45.8%; thus, R^2 and validation models can vary by 5% (40.8% 50.8%). Table X shows the split sample R^2 as 48.4% and 44.1%, whereas the full sample is 45.8%. The split sample validation's value is confirmed; i.e., cross-validation is verified and supports our finding's generalizability.

Table VIII, IX and X here

Discussion

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Appraising reliability and validity tests

Internal consistency is the degree to which various multidimensional construct measures correlate with the scale (Hair *et al.,* 2003). They indicate the item's homogeneity in the measure that captures the construct after administering the instrument. All approaches adopted for testing reliability and validity are as shown in Table XI.

Table XI here

Adopting Cronbach's alpha and split half technique show that the scale demonstrates reliability. The instrument was piloted and subjected to expert evaluation; i.e., demonstrating content validity. Concurrent and predictive validity have demonstrated criterion related validity. Discriminant validity was not assessed, convergent validity demonstrated construct validity. Aagja and Garg (2010); Kaul (2007) did not successfully demonstrate discriminant validity. Although there are different studies that measured healthcare service quality, only Duggirala *et al.,* (2008) tested and supported all three: content, criterion and construct validity (Table XII). Pai and Chary (2013) observed that hospital service quality studies have neither evaluated all three validity methods nor discussed them. All three validity methods applied to Pai and Chary's (2016) framework demonstrate instrument validity. The instrument, therefore, is deemed both reliable and valid for studying hospital service quality.

Table XII here

Pai and Chary's (2016) conceptual framework, comprising nine dimensions, measures hospital service quality using multiple regression. Results indicate that the null hypothesis can be rejected. According to Salkind (2000), an R (correlation coefficient) between 0.4 to 0.6 is moderate; 0.6 to 0.8 is strong and 0.8 to 1 is a strong to perfect association. The R in our study is 0.677 (Table VII) indicating a strong association between dependent and independent variables ($p < 0.001$). The model is also cross validated using adjusted R^2 and data splitting (Field, 2013).

Conclusion

Our objective was twofold, first to evaluate the Pai and Chary (2016) conceptual framework's reliability and validity and second to evaluate their proposed nine-dimension structure for measuring patient-perceived hospital service quality. Our study provides empirical evidence that the scale developed to measure hospital service quality is reliable and valid. We show that the nine dimensions predict hospital service quality.

Theoretical contribution and managerial implications

The study has strong implications for hospital quality managers seeking to measure hospital service quality who can implement the structured questionnaire in any hospital setting. Managers can also use the instrument to compare service quality among different combinations such as teaching, public and corporate hospitals.

Scope for future studies

Our study is cross sectional and limited to Karnataka state hospitals. Future studies should use a longitudinal study to provide greater diagnostic value. The instrument's validity and reliability can also be tested in other developing countries.

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

References

Aagja, J.P. and Garg, R. (2010), 'Measuring perceived service quality for public hospitals (PubHosQual) in the Indian context', *International Journal of Pharmaceutical and Healthcare Marketing*, Vol.4 No.1, pp.60-83.

Akter, S.M., Upal, M. and Hani, U. (2008), 'Service quality perception and satisfaction: a study over suburban public hospitals in Bangladesh', *Journal of Service Research*, February Special Issue, pp.125-146.

Al-Qatari, G. and Haran, D. (1999), 'Determinants of users' satisfaction with primary health care settings and services in Saudi Arabia', *International Journal for Quality in Health Care*, Vol. 11 No.6, pp.523-31.

Amira, E. (2008), 'Patient satisfaction in Japan', *International Journal of Health Care Quality Assurance*, Vol. 21 No.7, pp.692-705.

Andaleeb, S.S. (2000), 'Public and private hospitals in Bangladesh: service quality and predictors of hospital choice', *Health Policy and Planning*, Vol. 15 No. 1, pp.95-102.

Andres, L. (2012), *Designing and doing survey research*, Sage Publishers India Private Ltd., New Delhi.

Arasli, H., Ekiz, E.H. and Katircioglu, S.T. (2008), 'Gearing service quality into public and private hospitals in small islands empirical evidence from Cyprus', *International Journal of Health Care Quality Assurance*, Vol.21 No.1, pp.8-23.

Ashok, A.D. (2014). Telangana: Did you know how India got its 29 states and 7 UTs?', <http://indiatoday.intoday.in>, accessed June 2016.

Bahia, K. and Nantel, J. (2000), 'A reliable and valid measurement scale for the perceived service quality of banks', *International Journal of Bank Marketing*, Vol.18 No.2, pp.84-91.

Bloom, P.N. and Reeve, T. (1990), 'Transmitting signals to consumers for competitive advantage', *Business Horizons*, Vol. 33 No.4, pp.58-66.

Boshoff, C. and Terblanche, N.S. (1997), 'Measuring retail service quality: a replication study', *South African Journal of Business Management*, Vol.28 No.4, pp.123-128.

Brady, M.K. and Cronin, Jr. J. J. (2001), 'Some new thoughts on conceptualizing perceived service quality: a hierarchical approach', *Journal of Marketing,* Vol.65 No.3, pp.34-49.

Camilleri, D. and O'Callaghan, M. (1998), 'Comparing public and private hospital care service quality', *International journal of Health Care Quality Assurance*, Vol.11 No 4, pp. 127-133.

Campbell, D.T. and Fiske, D.W. (1959), 'Convergent and discriminant validation by the multitrait- multimethod matrix', *Psychological Bulletin*, Vol.56 No.2, pp. 81-105.

Candell, G. and Hulin, H. (1987), 'Cross-language and cross-cultural comparisons in scale translation', *Journal of Cross-Cultural Psychology*, Vol.17 No. 3, pp.417-440.

Carman, J.M. (1990), 'Consumer perceptions of service quality: an assessment of the SERVQUAL dimensions', *Journal of Retailing*, Vol.66 No.1, pp.33-55.

Choi, K.S., Cho, W.H., Lee, S., Lee, H. and Kim, C. (2004), 'The relationships among quality, value, satisfaction and behavioral intention in health care provider choice: A South Korean study', *Journal of Business Research*, Vol. 57 No.8, pp.913-921.

Clark, H.H. and Schober, M.F. (1992), 'Asking questions and influencing answers', in Tanur, J.M. (Eds.), *Questions about Questions: Inquires into the cognitive bases of surveys*, Russell Sage Foundation, New York, pp. 15-48.

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Collins, D. (2003), 'Pretesting survey instruments: an overview of cognitive methods', *Quality of Life Research*, Vol.12 No.3, pp.229-238.

Cooper, D.R., Schindler, P.S. and Sharma, J.K. (2012), *Business Research Methods*, 11th ed., Tata McGraw Hill Publishing Company Limited, New Delhi.

Crompton, J.L. and Love, L.L. (1995), 'The predictive validity of alternative approaches of evaluating quality of a festival', *Journal of Travel Research,* Vol.34 No.2, pp.123-138.

Czaja, R. (1998), 'Questionnaire pretesting comes of age', *Market Bull*, Vol.9, pp.52-66.

Dagger, T.S., Sweeney, J.C. and Johson, L.W. (2007), 'A hierarchical model of health service Quality', *Journal of Service Research*, Vol. 10 No. 2, pp. 123-42.

DeVellis, R.F. (1991), *Scale Development: Theory and Applications*, Sage Publications, California.

Duggirala, D., Rajendran, C. and Anantharaman, R.N. (2008), 'Patient perceived dimensions of total quality service in healthcare', *Benchmarking: An International Journal*, Vol. 15 No.5, pp. 560-583.

Easterby-Smith, M., Thorpe, R. and Lowe, A. (2002), *Management Research: An Introduction*, 2nd ed., Sage Publication India Private Limited, New Delhi.

Field, A. (2013), *Discovering Statistics using IBM SPSS Statistics*, 4th ed., Sage Publications Private Ltd., New Delhi.

Finn, A. and Kayande, U. (2004), 'Scale modification: alternative approaches and their consequences', *Journal of Retailing*, Vol.80 No.1, pp.37-52.

Forsyth, B.H. and Lessler, J.T. (1991), 'Cognitive laboratory methods: a taxonomy', in Biemer, P.P., Groves, R.M., Lyberg, L.E. Mathiowetz, N.A. and Sudman, S. (Eds), *Measurement Errors in Surveys*, Wiley, New York, pp.393-418.

Greenfield, T.K. and Attkisson, C.A.C. (1989), 'Steps Toward a Multifactorial Satisfaction Scale for Primary Care and Mental Health Services', *Evaluation and Program Planning*, Vol. 12 No.3, pp.271-279.

Hair, J.F. Jr, Bush, R.P. and Ortinau, D.J. (2003), *Marketing Research: Within a Changing Information Environment*, 2nd ed., Tata McGraw Hill, New Delhi.

Hansen, P.M., Peters, D.H., Vishwanathan, K., Rao, K.D., Mashkoor, A. and Burnham, G. (2008), 'Client perceptions of the quality of primary care services in Afghanistan', *International Journal for Quality in Health Care*, Vol.20 No.6, pp. 384-391.

Headley, D.E. and Miller, S.J. (1993), 'Measuring service quality and its relationship to future consumer behaviour', *Journal of Health Care Marketing*, Vol. 13 No. 4, pp. 32-40.

Hertzog, M.A. (2008), 'Considerations in determining sample size for pilot studies', *Research in Nursing and Health,* Vol.31 No.2, pp. 180-91.

Ho, R. (2006), *Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS*, Taylor and Francis group, USA.

Jabnoun, N. and Chakar, M. (2003), 'Comparing the quality of private and public hospitals', *Managing Service Quality*, Vol.13 No.4, pp.290-299.

Jobe, J.B. and Mingay, D.J. (1991), 'Cognition and Survey measurement: History and overview', *Applied Cognitive Psychology*, Vol.5 No.3, pp.175-192.

Kara, A., Lonial, S., Tarim, M. and Zaim, S. (2005), 'A paradox of service quality in Turkey', *European Business Review*, Vol. 17 No. 1, pp. 5-20.

Kaul, S. (2007), 'Measuring retail service quality: examining applicability of international research perspectives in India', *Vikalpa*, Vol.32 No.1, pp.15-26.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Kirchhoff, K.T. (1999), 'Design of questionnaires and structured interviews', in Mateo, M.A. and Kirchhoff, K.T. (Eds), *Using and conducting nursing research in the clinical setting*, 2nd ed., W.B. Saunders, Philadelphia, pp.229-239.

Lackey, N.R. and Wingate, A.L. (1998), 'The pilot study: one key to research success', in Brink, P.J. and Wood, M.J. (Eds), *Advanced Design in Nursing Research*, 2nd ed., Sage Publications, California, pp.375-386.

Likert, R. (1932), 'A technique for the measurement of attitudes', *Archives of Psychology*, Vol.140, pp.44-53.

Lohr, S. (2015), *Dataism: Inside the big data revolution*, Oneworld Publications, London.

Malhotra, N.K. and Dash, S. (2012), *Marketing research an applied orientation*, 6th ed., Pearson Education Inc., New Delhi.

Malinowski, B. (1935), *Coral Gardens and Their Magic, 2: The Language of Magic and Gardening*, Allen & Unwin, Sydney.

McGorr, S. (2000), 'Measurement in a cross-cultural environment: survey translation issues', *Qualitative Market Research*, Vol. 3 No. 1, pp.74-81.

Mostafa, M.M. (2005), 'An empirical study of patients' expectations and satisfactions in Egyptian hospitals', *International Journal of Health Care Quality Assurance*, Vol.18 No.6, pp.516-532.

Nunnally, J.C. (1978), *Psychometric Theory*, 2nd ed., McGraw-Hill, New York.

Oliver, R. L. (1997), Satisfaction: A Behavioral Perspective on the Consumer, McGraw-Hill, New York.

Pai, Y.P and Chary, S.T. (2013), 'Hospital service-quality dimensions: a critical review', *International Journal of Health Care Quality Assurance*, Vol.26 No.4, pp. 308 – 340.

Pai, Y.P and Chary, S.T. (2016), 'Measuring patient-perceived hospital service quality: a conceptual framework', *International Journal of Health Care Quality Assurance*, Vol.29 No.3, pp. 300 – 323.

Parasuraman, A., Grewal, D. and Krishnan, R. (2004), Marketing Research, 1st Indian adaptation, Biztantra, New Delhi.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1985), 'A conceptual model of service quality and its implications for future research', *Journal of Marketing*, Vol.49 No.4, pp.41-50.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988), 'SERVQUAL: A Multiple-item scale for measuring consumer perceptions of service quality', *Journal of Retailing*, Vol. 64 No.1, pp.12-40.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1994), 'Alternative scales for measuring service quality: a comparative assessment based on psychometric and diagnostic criteria', *Journal of Retailing*, Vol. 70 No.3, pp.201-230.

Pedhazur, E.J. (1997), 'A cross-validation approach to sample size determination', *Journal of the American Statistical Association*, Vol.69, pp.214-218.

Ping, R.A. (2004) 'On assuring valid measures for theoretical models using survey data', *Journal of Business Research*, Vol.57 No.2, pp. 125-141.

Rao, K.D., Peters, D.H. and Bandeen-Roche, K. (2006), 'Towards patient centered health services in India—a scale to measure patient perceptions of quality', *International Journal for Quality in Health Care*, Vol.18 No.6, pp.414-421.

Ramsaran-Fowdar, R. (2008), 'The relative importance of service dimensions in a healthcare setting', *International Journal of Health Care Quality Assurance*, Vol. 21 No.1, pp. 104-124.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Reynoso, J. and Moore, B. (1995), 'Towards the measurement of internal service quality', *International Journal of Service Industry Management*, Vol.6 No.3, pp.64-83.

Salkind, N. J. (2000), *Statistics for people who think they hate statistics*, Sage Publication, California.

Schwarz, N. (1997), 'Questionnaire design: The rocky road from concepts to answers', in Lyberg, L., Biemer, P., Collins, M., de Leeuw, E., Dippo, C., Schwarz, N. and Trewin, D. (Eds.), *Survey Measurement and Process Quality*, Wiley, New York, pp.29-45.

Sekaran, U. and Bougie, R. (2010), *Research methods for business: A skill building approach*, John Wiley & Sons, Delhi.

Sower, V., Duffy, J., Kilbourne, W., Kohers, G. and Jones, P. (2001), 'The dimensions of service quality for hospitals: development and use of the KQCAH scale', Health Care Management Review, Vol. 26 No. 2, pp. 47-59.

Stevens, J.P. (2002), *Applied Multivariate Statistics for the Social Sciences*, 4th ed., Erlbaum, Hillsdale, New Jersey.

Sudman, S. and Bradburn, N. (1987), *Asking Questions*, Jossey-Bass, San Francisco.

Taner, T. and Antony, J. (2006), 'Comparing public and private hospital care service quality in Turkey', *Leadership in Health Services*, Vol.19 No.2, pp. i-x.

Tabachnick, B. G., and Fidell, L. S. (2014), *Using Multivariate Statistics*, 6th ed., Pearson Education Limited, Harlow.

Taylor, S. A. and Baker, T.L. (1994), 'An Assessment of the Relationship between Service Quality and Customer Satisfaction in the Formation of Consumers' Purchase Intentions', *Journal of Retailing*, Vol.70 No.2, pp.163-178.

Tourangeau, R., Rips, L.J. and Rasinski, K. (2000), *The Psychology of Survey Response*, Cambridge University Press, Cambridge.

Wong, G.K.M., Lu, Y. and Yuan, L.L. (2001), 'SCATTR: an instrument for measuring shopping centre attractiveness', *International Journal of Retail and Distribution Management*, Vol.29 No.2, pp.76-86.

Zeithaml, V.A., Berry, L.L. and Parasuraman, A. (1996), 'The Behavioral Consequences of Service Quality', *Journal of Marketing*, Vol. 60 (April), pp. 31-46.

Table I: Demographic profile

Demographic Factors	Valid Items	Frequency	Percentage
Gender	Male	267	44.4
	Female	327	54.3
Age (in Years)	18-27	163	27.1
	28-37	159	26.4
	38-47	102	16.9
	48-57	68	11.2
	58-67	64	10.6
	68 and above	37	6.1

Table II: Cronbach alpha values

Constructs	Cronbach's Alpha

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Healthscape	0.875
Personnel	0.869
Hospital Image	0.739
Trustworthiness	0.805
Clinical Care Process	0.823
Communication	0.869
Relationship	0.845
Personalization	0.730
Administrative Procedures	0.896
Perceived Service Quality	0.873
Customer Satisfaction	0.811
Behavioural Intentions	0.882

Table III: Bivariate correlations among the dimensions

	HS	PL	HI	TW	CCP	COM	REL	PER	AP	PSQ	SAT	BI
HS	1											
PL	.668	1										
HI	.509	.574	1									
TW	.603	.702	.508	1								
CCP	.600	.675	.604	.631	1							
COM	.605	.673	.505	.646	.730	1						
REL	.420	.462	.386	.454	.472	.567	1					
PER	.446	.494	.423	.418	.523	.556	.582	1				
AP	.526	.565	.450	.586	.581	.693	.520	.521	1			
PSQ	.55	.551	.483	.548	.564	.587	.445	.414	.531	1		
SAT	.532	.604	.546	.585	.617	.591	.412	.451	.528	.633	1	
BI	.518	.575	.534	.527	.587	.562	.384	.441	.462	.651	.765	1

HS - Healthscape, PL - Personnel, HI - Hospital Image, TW- Trustworthiness, CCP - Clinical Care Process, COM - Communication, REL - Relationship, PER – Personalization, AP - Administrative Procedures, PSQ – Perceived Service Quality, SAT – Satisfaction,
BI – Behavioural Intentions.
All are positive and significant at 0.001.

Table IV: One-way ANOVA for assessing convergent validity – independent variable service quality

Descriptive Variables	Degrees of freedom	F	Sig.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

	Between Groups	Within Groups		
Healthscape	13	588	24.219	.001
Personnel	13	588	28.381	.001
Image	13	588	23.463	.001
Trustworthiness	13	588	28.694	.001
Clinical Care Process	13	588	29.622	.001
Communication	13	588	31.435	.001
Relationship	13	588	17.029	.001
Personalization	13	588	13.205	.001
Administrative Procedures	13	588	25.967	.001

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Table V: Validity tests

Validity	Measures used	Correlation coefficient (r)
Convergent validity	Perceived service quality	0.665 ($p<0.001$)
Predictive validity	Repurchase intentions	0.621 ($p<0.001$)

Table VI: Collinearity statistics

Independent Variable	Tolerance*	VIF*
Healthscape (HS)	.475	2.104
Personnel (PL)	.347	2.879
Hospital Image (HI)	.567	1.765
Trustworthiness (TW)	.413	2.421
Clinical Care Process (CCP)	.353	2.831
Communication (COM)	.308	3.248
Relationship (REL)	.559	1.787
Personalization (PER)	.544	1.838
Administrative Procedures (AP)	.455	2.198

*assessed using Enter regression method

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Table VII: Model summary

R	R Square	Adjusted R Square	F	Sig.	Durbin-Watson
.677	.458	.451	71.65	.001	1.987

Table VIII: Split data model summary - estimation sample – 0

R		R Square	Adjusted R Square	F	Sig.	Durbin-Watson	
Split = 0.00 (Selected)	split ~= 0.00 (Unselected)					Split = 0.00 (Selected)	split ~= 0.00 (Unselected)
.696	.623	.484	.478	75.059	.001	2.000	2.053

Table IX: Split data model summary - validation sample – 1

R		R Square	Adjusted R Square	F	Sig.	Durbin-Watson	
Split = 1.00 (Selected)	split = 1.00 (Unselected)					Split = 1.00 (Selected)	split ~= 1.00 (Unselected)
.664	.652	.441	.433	53.618	.001	2.056	1.944

Table X: Summary - Validation results: stepwise regression

	Entire data	Estimation sample	Validation sample
ANOVA significance (sig<=0.05)	<0.001	<0.001	<0.001
R^2	0.458	0.484	0.441

Table XI: Reliability and validity tests

Reliability	Content Validity	Criterion related Validity	Construct Validity
split-half reliability tests	Literature	Predictive validity (2 tests)	Convergent validity (2 tests)
inter-item consistency reliability	Pilot testing	Concurrent validity (2 tests)	

Table XII: Reliability and validity a comparison with established scales

Sl. No.	Particulars	Parasuraman *et al.,* (1988)	Aagja and Garg (2010)	Duggirala *et al.,* (2008)	Pai and Chary (2016)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

1	Instrument	SERVQUAL	PubHosQual		
2	Dimensional Structure	Five	Five	Seven	Nine
3	Application Domain	All services	Public Hospitals	Hospital Services	Hospital Services
4	Reliability	$0.72 < \alpha < 0.86$	$0.58 < \alpha < 0.89$	$0.775 < \alpha < 0.906$	$0.73 < \alpha < 0.896$
5	Content Validity	Measured	Measured	Measured	Measured
6	Construct Validity	Convergent Validity	Convergent and nomological validity are successful; discriminant validity had mixed results.	Convergent Validity	Convergent Validity
7	Criterion related validity			Concurrent validity	Both concurrent and predictive validity are measured

Appendix 1: Questionnaire

1. Modern and up-to-date equipment (e.g., Computerized tomography (CT) and Magnetic Resonance Imaging (MRI), patient information and billing) to serve patients more effectively.
2. Physical facilities are visually appealing.
3. Facility adequacy (e.g., wards, beds, operation theatre, Intensive Care Unit (ICU)).
4. Cleanliness (toilets, rooms and wards).
5. Infection-free environment/treatment.
6. Hospital staff follow hygienic care and procedures (e.g., wearing gloves).
7. Employees are dressed neatly.
8. Drugs are available.
9. Comfortable ambient conditions with proper lighting.
10. The hospital has an appealing atmosphere.
11. The hospital has clean rooms without foul smell.
12. There are sufficient waiting areas for patients and families.
13. It is easy to find my way in the hospital.
14. It is easy to find care facilities (laboratory, Doctor's office).
15. It is easy to use the amenities (public telephone, cafeteria, etc.).
16. Courtesy is shown by hospital staff towards patient and patient party.
17. Doctors and nurses are available as and when required.
18. The doctors are competent and skilful.
19. Knowledgeable nurses.
20. Paramedical and support staff are competent.
21. The doctors are friendly, caring and understand patient's feelings and needs.
22. Doctors talked to me frankly and politely.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

23. Nurses give prompt and timely attention.
24. Nursing staff are polite and well-mannered.
25. Staff at the hospital are polite.
26. Doctors are professional.
27. Good doctors are available in the hospital.
28. The hospital has positive reputation.
29. Sincerity, honesty and ethics followed in providing medical services.
30. The hospital staff run various programs for patients to support different societal sections.
31. The hospital provides medical services with nominal cost to the needy patients.
32. Patient privacy and confidentiality are maintained by hospital staff.
33. Hospital staff provide services as promised and on time.
34. Equal treatment for all.
35. Confidence in the doctor who treated you in the hospital.
36. The hospital provides patients with services beyond medical treatment.
37. Faultless assessment of health conditions by doctors.
38. Explanations are provided by the doctor about health status, medical tests.
39. Treatment procedures and outcomes are explained.
40. Medical advice and instructions provided by doctor.
41. Diagnosis is only made after careful examination.
42. Doctors spent enough time examining the patient.
43. Doctors provide information quickly.
44. My family was told what they needed to know.
45. Hospital staff provided adequate information about my illness/treatment(s).
46. Information can be easily obtained.
47. Obtaining information from hospital administrative personnel (e.g., admission, treatment, discharge) is easy.
48. Extent to which doctors' answer patient's questions and explain treatment that I could understand.
49. I feel good about the interaction I have with doctors at the hospital.
50. I feel good about the interaction I have with nurses at the hospital.
51. I feel good about the interaction I have with staff at the hospital.
52. I have built a close relationship with some staff at the hospital.
53. I have built a close relationship with the doctor at the hospital.
54. I have built a close relationship with nurses at the hospital.
55. I always get personalized attention from staff at the hospital.
56. Hospital staff treat me as a human being and not just a patient.
57. The doctor calls my name while addressing me.
58. Waiting time to consult the doctors is minimum.
59. Staff provide right patient services the first time, every time.
60. Reasonable waiting time spent for diagnostic test and treatments.
61. Time between the two successive processes is minimum.
62. The process for setting up the appointment was simple and easy.
63. Appointments at the hospital run on time.
64. The hospital's records and documentation are error free.
65. The interaction among department staff is well managed.
66. I believe the hospital is well-managed.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Perceived Service Quality (Brady and Cronin, 2001; Dagger et al., 2007; Parasuraman et al., 1988)

67. The overall service quality provided by hospital staff is excellent.
68. The overall service quality provided by hospital staff is impressive.
69. The service provided by the hospital is a high standard.
70. I believe the hospital offers service that is superior in every way.

Satisfaction (Dagger et al., 2007; Greenfield and Attkisson 1989; Oliver 1997)

71. My feelings towards the hospital are very positive.
72. I feel good about coming to this hospital for my treatment.
73. Overall, I am satisfied with the hospital and the service it provides.
74. I feel satisfied that the treatment results are the best that can be achieved.
75. The extent to which my treatment has produced the best possible outcome is satisfying.

Behavioral Intention (Dagger et al., 2007; Headley and Miller 1993; Taylor and Baker 1994; Zeithaml et al., 1996)

76. If I had to start treatment again, then I would want to come to this hospital.
77. I would highly recommend the hospital to other patients.
78. I have said positive things about the hospital to my family and friends.
79. I intend to continue having treatment, or any follow-up care I need, at this hospital.
80. I have no desire to change hospital.
81. I intend to follow the medical advice given to me at the hospital.
82. I am glad to have my treatment at this hospital rather than somewhere else.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Patient-perceived hospital service quality: an empirical assessment

Introduction

Healthcare is considered a 'credence' good – an offering that consumers will never be able to evaluate owing to deficient medical knowledge (Bloom and Reeve, 1990). Additionally, conceptualising and measuring service quality in a healthcare setting is more important and simultaneously more complex (Taner and Antony, 2006). However, researchers need to come up with ways to measure healthcare service quality, because unless we measure we cannot manage and improve healthcare services (Lohr, 2015). The literature indicates that there is a variability and confusion in how quality is conceptualized and operationalized (Sower *et al.,* 2001). Many researchers attempt to define and conceptualize hospital service quality. Pai and Chary (2013), for example, show that the Parasuraman *et al.,* (1985, 1988) SERVQUAL/modified SERVQUAL questionnaire is conventionally practiced in healthcare. They also pointed out studies where SERVQUAL items have not loaded onto respective dimensions highlighting a five-component structure that is lacking, suggesting a new questionnaire needs developing. New instruments, designed for healthcare setting: PRIVHEALTHQUAL (Ramsaran-Fowdar, 2008) for private and PubHosQual (Aagja and Garg, 2010) for public settings, have emerged. Studies adopting these instruments are scarce because these scales are hospital specific and not a general scale that measures hospital service quality in any hospital context. An instrument that could measure hospital service quality (HSQ) in any hospital settings has gained importance. Pai and Chary (2016) proposed their conceptual framework for measuring hospital service quality using nine dimensions. Their framework, unlike other instruments such as SERVQUAL (Parasuraman *et al.,* 1988), PRIVHEALTHQUAL (Ramsaran-Fowdar, 2008) and PubHosQual (Aagja and Garg, 2010) has an additional dimension (relationship) in line with researchers such as Carman (1990) and Reynoso and Moore (1995), who suggested adding service specific dimensions to SERVQUAL, there exists no empirical studies to support or refute their conceptual framework. As there are no research studies that used Pai and Chary's framework (2016), our purpose is to empirically appraise this framework and addresses literature gaps.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Objectives

We conducted our study with the objective to test the Pai and Chary (2016) conceptual framework's validity and reliability. We also aim to assess HSQ through nine dimensions: (i) healthscape; (ii) personnel; (iii) hospital image; (iv) trustworthiness; (v) clinical process; (vi) communication; (vii) relationship; (viii) personalization and (ix) administrative procedures.

Structured questionnaire

The questionnaire included:

1. Respondent demographics – gender, age, occupation, educational level, marital status and income.
2. Sixty-six items covering nine dimensions (Pai and Chary, 2016).
3. Four items measuring respondents' hospital service quality perception.
4. Twelve items that measured respondent satisfaction and behaviour.

A Likert (1932) scale, having a balanced rating with equal categories above and below the midpoint, was used with each item; anchored with verbal statements like 'strongly disagree', 'disagree', 'neither agree or disagree', 'agree', 'strongly agree'. All items were phrased

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

positively (Easterby-Smith *et al.,* 2002; Pai and Chary, 2013; Parasuraman *et al.,* 1994). Sixty-six hospital service quality items were categorized into nine dimensions: (i) healthscape (15 items); (ii) personnel (11); (iii) hospital image (5); (iv) trustworthiness (5); (v) clinical process (6); (vi) communication (9); (vii) relationship (3); (viii) personalization (3); and (ix) administrative procedures (9 items). Four perceived service quality, twelve satisfaction and behavioural intention items were included (Appendix 1).

Piloting

Piloting a new instrument is imperative (DeVellis, 1991; Kirchhoff, 1999). Before the pilot test is performed, it is advisable to obtain the original evaluation (DeVellis, 1991). Andres (2012) recommends that conducting a pilot study by involving colleagues, friends and family members who assume an audience role. Our questionnaire was subjected to the pre-testing process by circulating it among ten colleagues, friends and family members. In any research, respondents often misunderstand words or concepts. Although communication difficulties exist, respondents still provide legitimate answers to survey questions (Clark and Schober, 1992; Tourangeau *et al.,* 2000). To overcome these problems, Collins (2003, p.230) suggested cognitively testing survey questions that helps 'to identify how and where the question fails to achieve its measurement purpose'. Although there are various cognitive methods that have been developed and applied to instrument testing, such as cognitive interviewing, paraphrasing, card sorts, vignettes, confidence ratings and response latency timing (Czaja, 1998; Forsyth and Lessler, 1991; Jobe and Mingay, 1991); cognitive interviewing is becoming widespread (Schwarz, 1997). Cognitive interviewing (involving two methods: think aloud interviewing and probing) and paraphrasing were used in testing. On completion, the instrument was tried in a hospital with patients as respondents. According to Lackey and Wingate (1998), pilot testing a newly developed instrument should be undertaken with respondents selected from the same population from which the subjects in the major study will be selected. Consequently, the questionnaire was tested with respondents from teaching, corporate and public hospitals. Pre-testing ensured correct phrasing, format, length and question sequence. Pretesting was done with 30 respondents using Hertzog's (2008) guidance. The questionnaire was corrected after feedback.

Data collection

Data were collected in three hospitals – teaching, corporate and public in Karnataka. India has 29 states and seven union territories (Ashok, 2014), and every state has a specific language. Although India is an English-speaking country, Kannada is Karnataka's local language. Hence, the questionnaires were in Kannada and translated into the Malayalam language as few hospitals had Malayalam speaking patients in significant numbers owing to hospital's proximity to another state (Kerala), where the official language is Malayalam. Consequently, questionnaires were administered in three languages: English; Kannada and Malayalam. There are several studies in which the questionnaire was administered in different languages specific to the country studied; e.g., English and Maltese (Camilleri and O'Callaghan, 1998); English and Arabic (Jabnoun and Chakar, 2003); English and Turkish (Kara *et al.,* 2005); English and Gujrati (Aagja and Garg, 2010); and English and Bengali (Akter *et al.,* 2008). The Kannada/Malayalam versions were created through careful translation and back-translation techniques (Candell and Hulin, 1987; McGorr, 2000). First, the questionnaire was translated into Kannada/Malayalam and then the Kannada/Malayalam items were back-translated into English by a bilingual expert to ensure that

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

the original content was kept intact. In translating the scale items into Kannada/Malayalam, Malinowski's (1935) four-step translation technique was implemented:

1. An interlinear, or word-by-word, translation.
2. A 'free' translation in which clarifying terms, conjunctions, etc., are added and words reinterpreted.
3. Analysing and collating two translations leading to:
4. A contextual specification of meaning.

An attempt was made to remove discrepancies between English and Kannada/Malayalam questionnaire versions. No individual items were found to be problematic in translation. Previous studies reported backward translation; e.g., English to Bengali (Akter *et al.,* 2008; Andaleeb, 2000); English to Japanese (Amira, 2008); English to Arabic (Mostafa, 2005); English to Hindi (Rao *et al.,* 2006). Accordingly, our instrument was translated from English to Kannada and Malayalam languages.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Sample

The study population was defined as all patients 18 years or older with a stable mental and clinical condition during the data collection procedure. A stratified random sampling was adopted for the study. As the name implies, there is a stratification or segregation process followed by randomly selecting subjects from each stratum. The population is first divided into mutually exclusive groups which are relevant, appropriate and meaningful in the study context. The teaching, corporate and public hospitals constitute three strata. Hospitals were selected using systematic sampling and proportionate stratified sampling. The sampling frame was prepared for three strata and ten per cent from each stratum: three hospitals each from teaching and public hospitals and four corporate hospitals, totalling ten hospitals. Respondent from ten hospitals were chosen randomly (every fifth element in the population), both inpatients and outpatients yielding 602 respondents. The respondent's demographic profile is depicted in the Table I, comprising males (44%). Their sample's age ranged from 18 to 67+ years.

Table I here

Reliability and validity

Instruments designed to measure a specific concept should measure what is set out to measure. Reliability and validity are the two main criteria for measuring how good measures are in any research instrument. Reliability indicates stability and consistency with which the instrument measures the concept, thus assessing a measure's goodness (Sekaran and Bougie, 2010). It is important to calculate scale reliability, which refers to the extent to which a scale can reproduce the same results in repeated trials (Hair *et al.,* 2003). Consistency can be examined through the inter-item consistency reliability and split-half-reliability tests (Sekaran and Bougie, 2010). The internal consistency reliability test is acceptable when the reliability coefficient exceeds Nunnally's (1978) 0.7 reliability criterion.

The questionnaire included 82 items (Appendix 1). For internal consistency, the most popular test is coefficient alpha (Cronbach). For Pai and Chary's (2016) 66 items, the value is 0.965 and 0.972 for the 82-item instrument. The higher the coefficients, the better the measuring instrument (Sekaran and Bougie, 2010). Another technique for testing internal consistency is the

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

split-half technique, where items constituting the instrument are divided into two halves and the resulting half scores are correlated (Malhotra and Dash, 2012). For the 66-item version, Cronbach's Alpha was 0.93 and 0.947, correlation between forms was 0.793, Spearman–Brown Coefficient for equal length and unequal length is 0.884, Guttman Split-Half Coefficient is 0.878. For the 82-item questionnaire, Cronbach's Alpha was 0.944 and 0.956, correlation between forms was 0.821, Spearman–Brown Coefficient for equal length and unequal length is 0.902, Guttman Split-Half Coefficient is 0.897. We can conclude, therefore, that the instrument has adequate reliability. Cronbach's alpha is a widely-used reliability coefficient that assesses the entire scale's internal consistency and averages all possible split-half coefficients resulting from different ways of splitting the scale items ranging from 0 to 1. According to Malhotra and Dash (2012), it is appropriate to use internal consistency reliability for each dimension, if several items are used to measure each dimension. In our study, patient-perceived hospital service quality is measured using different dimensions, each is measured by several items, hence we used Cronbach's alpha to measure internal consistency. Cronbach's alpha for the twelve constructs considered in the study was measured and values for each construct are 0.70 and above, indicating a strong reliability (Table II).

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Table II here

Although reliability is a necessary contributor to validity, it is not a sufficient condition for validity (Sekaran and Bougie, 2010). Therefore, validity becomes equally important. Validity is 'the extent to which a rating scale truly reflects the underlying variable that it attempts to measure' (Parasuraman *et al.,* 2004, p. 294). Content validity, criterion-related validity and construct validity are generally discussed in research articles.

Content validity

Content validity is the degree to which items adequately represents all relevant items under study (Cooper *et al.,* 2012), determined using judgment and panel evaluation. For an attitudinal scale, content validity is an overall criterion that can be assessed only though a researcher's subjective judgment (Parasuraman *et al.,* 2004). We exercised judgment through carefully defining and analysing conceptual and empirical frameworks through an extensive literature review. Judges can attest to content validity (Sekaran and Bougie, 2010), so the questionnaire was subjected to expert review by practitioners and academics. Content/face validity was confirmed; i.e., the proposed measurement instrument measures the intended construct. The items for the current study were chosen from the literature (Sekaran and Bougie, 2010). Scales were refined using a pilot study. Patients revealed they had no difficulty understanding the questionnaire items indicating and confirming face validity (Arasli *et al.,* 2008). All steps ensured that the instrument possesses face validity. Although implementing a scale based on only face validity claims has a far greater impact than item's random deletion through mere formal scale refinement methods (Finn and Kayande, 2004), nevertheless, other validity tests are also conducted.

Criterion related validity (Ping, 2004) emphasizes that, for new measures, criterion validity should be assessed, which is established when the measure differentiates individuals on a criterion it is expected to predict. Criterion related validity can be established by concurrent or predictive validity (Sekaran and Bougie, 2010). Predictive validity indicates the instrument's ability to differentiate among individuals with reference to a future criterion (Sekaran and Bougie, 2010). The researcher needs to ensure that the validity criterion used is valid and the

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

intended measure is judged on four qualities: relevance; freedom from bias; reliability; and availability (Cooper *et al.,* 2012). Criterion-related validity was established using predictive validity, adopting Jabnoun and Chakar's (2003) technique, who correlated their service quality dimensions with overall service quality. Individual item scores on nine dimensions (listed earlier) were summed to obtain overall scores for each respondent. Scores were then correlated with a summated perceived service quality, satisfaction and behavioural intentions scale. All items were captured and measured using Likert scales with a five-point response format. A higher score indicated a more favourable response (Table III). All correlations were positive and statistically significant at 0.001 level, which establishes predictive validity.

Table III here

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Concurrent validity was assessed by examining the association between perceived service quality and a related variable, patient satisfaction (Hansen *et al.,* 2008). Patient satisfaction is a distant but related construct to perceived service quality (Choi *et al.,* 2004; Rao *et al.,* 2006). Although there are differences between patient satisfaction and perceived service quality, they are related and a moderate correlation is expected (Al-Qatari and Haran, 1999). Crompton and Love (1995) stressed that the two constructs are likely to be positively correlated, but unlikely to be linear. The correlation between perceived service quality and patient satisfaction is 0.633 (Table III), which indicates concurrent validity. Concurrent validity assessment is conducted in line with Duggirala *et al.,* (2008) such that in the present context, perceived service quality (PSQ) is chosen as the nine-dimension's outcome and a bivariate correlation analysis was carried out among all constructs, which have significant positive correlations with the criterion chosen, namely, PSQ (range 0.414 to 0.587, $p < 0.001$). According to Sekaran and Bougie (2010), construct validity testifies how well the results obtained from a measure fit the theories around which the test is designed and assessed through convergent and discriminant validity. Convergent validity is the degree to which scores on one scale correlate with scores on other scales designed to assess the same construct (Cooper *et al.,* 2012). Convergent validity is also assessed by calculating one-way analysis of variance (ANOVA) (Aagja and Garg, 2010; Bahia and Nantel, 2000; Parasuraman *et al.,* 1988; Wong *et al.,* 2001). The association between healthscape, personnel, hospital image, trustworthiness, clinical care process, communication, relationship, personalization and administrative procedures and PSQ was significant at less than 1% (Table IV), which means that the group differences are significant; i.e., good convergent validity.

Table IV here

Alternate methods for establishing criterion and construct validity
Campbell and Fiske's (1959) procedure has been used as an alternate method for establishing criterion and construct validity, adopted in similar studies such as Boshoff and Terblanche, 1997; Kaul 2007. Convergent validity (Table V) is confirmed as there is a high correlation between scale items and PSQ (correlation coefficient, $r = 0.665$, $p<0.001$). To assess predictive validity, respondents were asked whether they intended to return to the same hospital (repurchase intentions). Results confirm predictive validity ($r = 0.621$, $p< 0.001$).

Table V here

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Empirically validating nine patient-perceived hospital service quality dimensions
We propose the following hypothesis:

H_0: The heathscape, personnel, hospital image, trustworthiness, clinical process, communication, relationship, personalization and administrative procedure domains do not predict hospital service quality.

To test our hypothesis, a multiple regression approach is adopted with the nine dimensions as independent variables and PSQ as the dependent variable. Variables were examined for multicollinearity and autocorrelation before the regression analysis was run. Multicollinearity exists when there is a strong correlation between two or more predictors in a regression model leading to a regression model that fits the data, but no variable had a significant impact in predicting the dependent variable. In SPSS, there are collinearity diagnostics that are variance inflation factor (VIF) and tolerance statistics. The VIF indicates whether the predictor has a strong linear relationship with other predictor(s), while the tolerance statistic reciprocates the variance inflation factor (VIF). We used the Durbin-Watson test to identify any serial correlations between errors (Table VI and VII).

Table VI here

The tolerance values need to be greater than 0.1 and the VIF values should be less than 10 (Ho, 2006). From the tolerance and VIF values shown in Table VI, multicollinearity does not exist in the predictor values. Hence, the nine independent variables influence on dependent variables is analysed. The regression model is formed using the statistical (stepwise) regression method taking all nine independent variables together and examining their combined effect on the dependent variable. In the statistical (stepwise) regression model, the predictor variables' order is solely based on statistical criteria and variables that correlate most strongly with the dependent variable will be given entry priority. Our regression models require no references to theoretical considerations, and are primarily used when the researcher is unsure about the independent variable's predictive power (Ho, 2006). This method is apt for the present study and therefore stepwise regression is adopted, which includes forward inclusion, backward elimination and stepwise solution (Ho, 2006; Malhotra and Dash, 2012). To identify an optimal regression equation, one needs to compute combinational solutions in which all possible combinations are examined (Malhotra and Dash, 2012). The optimal regression equation in our study is computed using all possible combinations. However, the stepwise regression method is preferable to the forward method, as they reduce Type II errors (i.e., missing a predictor that does in fact predict the outcome) (Field, 2013). Regression analysis results are shown in Table VII.

Table VII here

The F value tests how well the regression model fits the data (Ho, 2006). In case the probability associated with the F statistics is small, the hypothesis that R-square is equal to zero is rejected. In our case, F is 71.65, $p < 0.001$ (Table VII); i.e., the hypothesis that there is no linear relationship between the predictor and the dependent variables is rejected. The R statistic indicates that the correlation coefficient as 0.677 and the R-square value is 0.458. For our

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

sample, the predictor variables explained 45.8 per cent of the variance in the dependent variable. Hence, the null hypothesis H_o is rejected. In Table VII, Durbin-Watson's d = 1.987. Durbin-Watson can vary between 0 and 4; i.e., 2 implies that the residuals are uncorrelated. If the value is greater than 2, then it indicates a negative correlation between adjacent residuals; whereas a value below 2 indicates a positive correlation (Field, 2013). Thus, Durbin-Watson statistic tests serial correlations between errors, a metric that assesses another linear model assumption. Field (2013) suggests that the stepwise method should be avoided except for exploratory model building. He recommends cross-validating the model if the stepwise method is used, suggesting two cross-validation methods: adjusted R^2 and data splitting.

Cross-validating the model using adjusted R square (R^2*)*

The adjusted R^2 shows how well our model generalizes. Although the adjusted R^2 should be the same or very close to R^2, it differs generally. This difference is known as shrinkage, which = R^2 - Adjusted R^2. As per model summary (Table VII) shrinkage is equal to 0.458 - 0.451 = 0.007 (about 0.7 per cent); i.e., the model's shrinkage is small. Although there are no shrinkage guidelines in the literature, it is valuable cross-validation such that one can be confident in the equation's generalizability if they are similar (Pedhazur, 1997). According to Field (2013), shrinkage means that if the model were derived from the population rather than a sample, then it would account for approximately 0.7 per cent less variance in the outcome. Stein's formula helps us to calculate adjusted R^2 and tells us how well the model cross-validates (Stevens, 2002), which is computed as: Adjusted R^2 = 1-[(n-1/n-k-1) (n-2/n-k-2) (n+1/n)] (1- R^2) where R^2 is the unadjusted value, n is total participants and k is total predictors in the model. Substituting R^2 as 0.458 (Table VII), n = 602 and k=9 in the above formula, we get 0.441, which is computed as 1-[(602-1/602-9-1) (602-2/602-9-2) (602+1/602)] (1- 0.458), which = 0.441. This value is like the observed 0.458, indicating that cross-validity is acceptable (Field, 2013, p.312).

Cross-validating the model using data splitting

The other cross-validation method is the data splitting approach. Cross-validation is a validity test that examines whether the regression model continues to hold on comparable data not used in the original estimation (Malhotra and Dash, 2012). In a cross-validation procedure, we estimate the regression model for the entire data, which are split into two parts: estimation and validation samples. The estimation sample incudes 50-90% of the total sample. The regression model is estimated using the estimation sample only. This model is compared with the entire sample. The estimated model is applied to the data in the validation sample to predict the dependent variable. In case there is a large discrepancy between R^2 for the smaller and larger samples, this indicate over fitting and weak generalizability (Tabachnick and Fidell, 2014). In other words, if there is a significant difference between R^2 for the larger sample (estimation sample) and smaller sample (validation sample) then there is weak generalizability. Using SPSS, the step-wise regression on a random selection is adopted as suggested by Field (2013). Cases are randomly divided into two parts for cross-validation. Creating a variable that contains information about the estimation sample and validation samples using 0 and 1 to indicate the estimation sample and validation sample. We split the data into two samples: estimation and validation samples with 325 (53.9%) and 277 (46.1%) respondents/samples using Malhotra and Dash, (2012) guidance that estimation sample includes 50-90% of the total sample. Stepwise regression is conducted on the split data; i.e., estimation and validation sample, and the R^2 values are compared with the original model (Malhotra and Dash, 2012; Tabachnick and Fidell, 2014).

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

The overall relationship between the dependent variable (PSQ) and the independent variable (nine dimensions) must be statistically significant for both validation analyses. Table VIII and IX indicate that significance is achieved. The R^2 for each validation must be within ±5% (plus or minus 5%) of the model's R^2 using the full sample. The entire sample's R^2 is 45.8%; thus, R^2 and validation models can vary by 5% (40.8% 50.8%). Table X shows the split sample R^2 as 48.4% and 44.1%, whereas the full sample is 45.8%. The split sample validation's value is confirmed; i.e., cross-validation is verified and supports our finding's generalizability.

Table VIII, IX and X here

Discussion

Appraising reliability and validity tests

Internal consistency is the degree to which various multidimensional construct measures correlate with the scale (Hair *et al.,* 2003). They indicate the item's homogeneity in the measure that captures the construct after administering the instrument. All approaches adopted for testing reliability and validity are as shown in Table XI.

Table XI here

Adopting Cronbach's alpha and split half technique show that the scale demonstrates reliability. The instrument was piloted and subjected to expert evaluation; i.e., demonstrating content validity. Concurrent and predictive validity have demonstrated criterion related validity. Discriminant validity was not assessed, convergent validity demonstrated construct validity. Aagja and Garg (2010); Kaul (2007) did not successfully demonstrate discriminant validity. Although there are different studies that measured healthcare service quality, only Duggirala *et al.,* (2008) tested and supported all three: content, criterion and construct validity (Table XII). Pai and Chary (2013) observed that hospital service quality studies have neither evaluated all three validity methods nor discussed them. All three validity methods applied to Pai and Chary's (2016) framework demonstrate instrument validity. The instrument, therefore, is deemed both reliable and valid for studying hospital service quality.

Table XII here

Pai and Chary's (2016) conceptual framework, comprising nine dimensions, measures hospital service quality using multiple regression. Results indicate that the null hypothesis can be rejected. According to Salkind (2000), an R (correlation coefficient) between 0.4 to 0.6 is moderate; 0.6 to 0.8 is strong and 0.8 to 1 is a strong to perfect association. The R in our study is 0.677 (Table VII) indicating a strong association between dependent and independent variables ($p < 0.001$). The model is also cross validated using adjusted R^2 and data splitting (Field, 2013).

Conclusion

Our objective was twofold, first to evaluate the Pai and Chary (2016) conceptual framework's reliability and validity and second to evaluate their proposed nine-dimension structure for measuring patient-perceived hospital service quality. Our study provides empirical evidence that the scale developed to measure hospital service quality is reliable and valid. We show that the nine dimensions predict hospital service quality.

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

Theoretical contribution and managerial implications

The study has strong implications for hospital quality managers seeking to measure hospital service quality who can implement the structured questionnaire in any hospital setting. Managers can also use the instrument to compare service quality among different combinations such as teaching, public and corporate hospitals.

Scope for future studies

Our study is cross sectional and limited to Karnataka state hospitals. Future studies should use a longitudinal study to provide greater diagnostic value. The instrument's validity and reliability can also be tested in other developing countries.

References

Aagja, J.P. and Garg, R. (2010), 'Measuring perceived service quality for public hospitals (PubHosQual) in the Indian context', *International Journal of Pharmaceutical and Healthcare Marketing*, Vol.4 No.1, pp.60-83.

Akter, S.M., Upal, M. and Hani, U. (2008), 'Service quality perception and satisfaction: a study over suburban public hospitals in Bangladesh', *Journal of Service Research*, February Special Issue, pp.125-146.

Al-Qatari, G. and Haran, D. (1999), 'Determinants of users' satisfaction with primary health care settings and services in Saudi Arabia', *International Journal for Quality in Health Care*, Vol. 11 No.6, pp.523-31.

Amira, E. (2008), 'Patient satisfaction in Japan', *International Journal of Health Care Quality Assurance*, Vol. 21 No.7, pp.692-705.

Andaleeb, S.S. (2000), 'Public and private hospitals in Bangladesh: service quality and predictors of hospital choice', *Health Policy and Planning*, Vol. 15 No. 1, pp.95-102.

Andres, L. (2012), *Designing and doing survey research*, Sage Publishers India Private Ltd., New Delhi.

Arasli, H., Ekiz, E.H. and Katircioglu, S.T. (2008), 'Gearing service quality into public and private hospitals in small islands empirical evidence from Cyprus', *International Journal of Health Care Quality Assurance*, Vol.21 No.1, pp.8-23.

Ashok, A.D. (2014). Telangana: Did you know how India got its 29 states and 7 UTs?', <http://indiatoday.intoday.in>, accessed June 2016.

Bahia, K. and Nantel, J. (2000), 'A reliable and valid measurement scale for the perceived service quality of banks', *International Journal of Bank Marketing*, Vol.18 No.2, pp.84-91.

Bloom, P.N. and Reeve, T. (1990), 'Transmitting signals to consumers for competitive advantage', *Business Horizons*, Vol. 33 No.4, pp.58-66.

Boshoff, C. and Terblanche, N.S. (1997), 'Measuring retail service quality: a replication study', *South African Journal of Business Management*, Vol.28 No.4, pp.123-128.

Brady, M.K. and Cronin, Jr. J. J. (2001), 'Some new thoughts on conceptualizing perceived service quality: a hierarchical approach', *Journal of Marketing,* Vol.65 No.3, pp.34-49.

Camilleri, D. and O'Callaghan, M. (1998), 'Comparing public and private hospital care service quality', *International journal of Health Care Quality Assurance*, Vol.11 No 4, pp. 127-133.

© Emerald Publishing Limited

This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Campbell, D.T. and Fiske, D.W. (1959), 'Convergent and discriminant validation by the multitrait- multimethod matrix', *Psychological Bulletin*, Vol.56 No.2, pp. 81-105.
Candell, G. and Hulin, H. (1987), 'Cross-language and cross-cultural comparisons in scale translation', *Journal of Cross-Cultural Psychology*, Vol.17 No. 3, pp.417-440.
Carman, J.M. (1990), 'Consumer perceptions of service quality: an assessment of the SERVQUAL dimensions', *Journal of Retailing*, Vol.66 No.1, pp.33-55.
Choi, K.S., Cho, W.H., Lee, S., Lee, H. and Kim, C. (2004), 'The relationships among quality, value, satisfaction and behavioral intention in health care provider choice: A South Korean study', *Journal of Business Research*, Vol. 57 No.8, pp.913-921.
Clark, H.H. and Schober, M.F. (1992), 'Asking questions and influencing answers', in Tanur, J.M. (Eds.), *Questions about Questions: Inquires into the cognitive bases of surveys*, Russell Sage Foundation, New York, pp. 15-48.
Collins, D. (2003), 'Pretesting survey instruments: an overview of cognitive methods', *Quality of Life Research*, Vol.12 No.3, pp.229-238.
Cooper, D.R., Schindler, P.S. and Sharma, J.K. (2012), *Business Research Methods*, 11th ed., Tata McGraw Hill Publishing Company Limited, New Delhi.
Crompton, J.L. and Love, L.L. (1995), 'The predictive validity of alternative approaches of evaluating quality of a festival', *Journal of Travel Research,* Vol.34 No.2, pp.123-138.
Czaja, R. (1998), 'Questionnaire pretesting comes of age', *Market Bull*, Vol.9, pp.52-66.
Dagger, T.S., Sweeney, J.C. and Johson, L.W. (2007), 'A hierarchical model of health service Quality', *Journal of Service Research*, Vol. 10 No. 2, pp. 123-42.
DeVellis, R.F. (1991), *Scale Development: Theory and Applications*, Sage Publications, California.
Duggirala, D., Rajendran, C. and Anantharaman, R.N. (2008), 'Patient perceived dimensions of total quality service in healthcare', *Benchmarking: An International Journal*, Vol. 15 No.5, pp. 560-583.
Easterby-Smith, M., Thorpe, R. and Lowe, A. (2002), *Management Research: An Introduction*, 2nd ed., Sage Publication India Private Limited, New Delhi.
Field, A. (2013), *Discovering Statistics using IBM SPSS Statistics*, 4th ed., Sage Publications Private Ltd., New Delhi.
Finn, A. and Kayande, U. (2004), 'Scale modification: alternative approaches and their consequences', *Journal of Retailing*, Vol.80 No.1, pp.37-52.
Forsyth, B.H. and Lessler, J.T. (1991), 'Cognitive laboratory methods: a taxonomy', in Biemer, P.P., Groves, R.M., Lyberg, L.E. Mathiowetz, N.A. and Sudman, S. (Eds), *Measurement Errors in Surveys*, Wiley, New York, pp.393-418.
Greenfield, T.K. and Attkisson, C.A.C. (1989), 'Steps Toward a Multifactorial Satisfaction Scale for Primary Care and Mental Health Services', *Evaluation and Program Planning*, Vol. 12 No.3, pp.271-279.
Hair, J.F. Jr, Bush, R.P. and Ortinau, D.J. (2003), *Marketing Research: Within a Changing Information Environment*, 2nd ed., Tata McGraw Hill, New Delhi.
Hansen, P.M., Peters, D.H., Vishwanathan, K., Rao, K.D., Mashkoor, A. and Burnham, G. (2008), 'Client perceptions of the quality of primary care services in Afghanistan', *International Journal for Quality in Health Care*, Vol.20 No.6, pp. 384-391.
Headley, D.E. and Miller, S.J. (1993), 'Measuring service quality and its relationship to future consumer behaviour', *Journal of Health Care Marketing*, Vol. 13 No. 4, pp. 32-40.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Hertzog, M.A. (2008), 'Considerations in determining sample size for pilot studies', *Research in Nursing and Health,* Vol.31 No.2, pp. 180-91.
Ho, R. (2006), *Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS*, Taylor and Francis group, USA.
Jabnoun, N. and Chakar, M. (2003), 'Comparing the quality of private and public hospitals', *Managing Service Quality*, Vol.13 No.4, pp.290-299.
Jobe, J.B. and Mingay, D.J. (1991), 'Cognition and Survey measurement: History and overview', *Applied Cognitive Psychology*, Vol.5 No.3, pp.175-192.
Kara, A., Lonial, S., Tarim, M. and Zaim, S. (2005), 'A paradox of service quality in Turkey', *European Business Review*, Vol. 17 No. 1, pp. 5-20.
Kaul, S. (2007), 'Measuring retail service quality: examining applicability of international research perspectives in India', *Vikalpa*, Vol.32 No.1, pp.15-26.
Kirchhoff, K.T. (1999), 'Design of questionnaires and structured interviews', in Mateo, M.A. and Kirchhoff, K.T. (Eds), *Using and conducting nursing research in the clinical setting*, 2nd ed., W.B. Saunders, Philadelphia, pp.229-239.
Lackey, N.R. and Wingate, A.L. (1998), 'The pilot study: one key to research success', in Brink, P.J. and Wood, M.J. (Eds), *Advanced Design in Nursing Research*, 2nd ed., Sage Publications, California, pp.375-386.
Likert, R. (1932), 'A technique for the measurement of attitudes', *Archives of Psychology*, Vol.140, pp.44-53.
Lohr, S. (2015), *Dataism: Inside the big data revolution*, Oneworld Publications, London.
Malhotra, N.K. and Dash, S. (2012), *Marketing research an applied orientation*, 6th ed., Pearson Education Inc., New Delhi.
Malinowski, B. (1935), *Coral Gardens and Their Magic, 2: The Language of Magic and Gardening*, Allen & Unwin, Sydney.
McGorr, S. (2000), 'Measurement in a cross-cultural environment: survey translation issues', *Qualitative Market Research*, Vol. 3 No. 1, pp.74-81.
Mostafa, M.M. (2005), 'An empirical study of patients' expectations and satisfactions in Egyptian hospitals', *International Journal of Health Care Quality Assurance*, Vol.18 No.6, pp.516-532.
Nunnally, J.C. (1978), *Psychometric Theory*, 2nd ed., McGraw-Hill, New York.
Oliver, R. L. (1997), Satisfaction: A Behavioral Perspective on the Consumer, McGraw-Hill, New York.
Pai, Y.P and Chary, S.T. (2013), 'Hospital service-quality dimensions: a critical review', *International Journal of Health Care Quality Assurance*, Vol.26 No.4, pp. 308 – 340.
Pai, Y.P and Chary, S.T. (2016), 'Measuring patient-perceived hospital service quality: a conceptual framework', *International Journal of Health Care Quality Assurance*, Vol.29 No.3, pp. 300 – 323.
Parasuraman, A., Grewal, D. and Krishnan, R. (2004), Marketing Research, 1st Indian adaptation, Biztantra, New Delhi.
Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1985), 'A conceptual model of service quality and its implications for future research', *Journal of Marketing*, Vol.49 No.4, pp.41-50.
Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988), 'SERVQUAL: A Multiple-item scale for measuring consumer perceptions of service quality', *Journal of Retailing*, Vol. 64 No.1, pp.12-40.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1994), 'Alternative scales for measuring service quality: a comparative assessment based on psychometric and diagnostic criteria', *Journal of Retailing*, Vol. 70 No.3, pp.201-230.

Pedhazur, E.J. (1997), 'A cross-validation approach to sample size determination', *Journal of the American Statistical Association*, Vol.69, pp.214-218.

Ping, R.A. (2004) 'On assuring valid measures for theoretical models using survey data', *Journal of Business Research*, Vol.57 No.2, pp. 125-141.

Rao, K.D., Peters, D.H. and Bandeen-Roche, K. (2006), 'Towards patient centered health services in India—a scale to measure patient perceptions of quality', *International Journal for Quality in Health Care*, Vol.18 No.6, pp.414-421.

Ramsaran-Fowdar, R. (2008), 'The relative importance of service dimensions in a healthcare setting', *International Journal of Health Care Quality Assurance*, Vol. 21 No.1, pp. 104-124.

Reynoso, J. and Moore, B. (1995), 'Towards the measurement of internal service quality', *International Journal of Service Industry Management*, Vol.6 No.3, pp.64-83.

Salkind, N. J. (2000), *Statistics for people who think they hate statistics*, Sage Publication, California.

Schwarz, N. (1997), 'Questionnaire design: The rocky road from concepts to answers', in Lyberg, L., Biemer, P., Collins, M., de Leeuw, E., Dippo, C., Schwarz, N. and Trewin, D. (Eds.), *Survey Measurement and Process Quality*, Wiley, New York, pp.29-45.

Sekaran, U. and Bougie, R. (2010), *Research methods for business: A skill building approach*, John Wiley & Sons, Delhi.

Sower, V., Duffy, J., Kilbourne, W., Kohers, G. and Jones, P. (2001), 'The dimensions of service quality for hospitals: development and use of the KQCAH scale', Health Care Management Review, Vol. 26 No. 2, pp. 47-59.

Stevens, J.P. (2002), *Applied Multivariate Statistics for the Social Sciences*, 4th ed., Erlbaum, Hillsdale, New Jersey.

Sudman, S. and Bradburn, N. (1987), *Asking Questions*, Jossey-Bass, San Francisco.

Taner, T. and Antony, J. (2006), 'Comparing public and private hospital care service quality in Turkey', *Leadership in Health Services*, Vol.19 No.2, pp. i-x.

Tabachnick, B. G., and Fidell, L. S. (2014), *Using Multivariate Statistics*, 6th ed., Pearson Education Limited, Harlow.

Taylor, S. A. and Baker, T.L. (1994), 'An Assessment of the Relationship between Service Quality and Customer Satisfaction in the Formation of Consumers' Purchase Intentions', *Journal of Retailing*, Vol.70 No.2, pp.163-178.

Tourangeau, R., Rips, L.J. and Rasinski, K. (2000), *The Psychology of Survey Response*, Cambridge University Press, Cambridge.

Wong, G.K.M., Lu, Y. and Yuan, L.L. (2001), 'SCATTR: an instrument for measuring shopping centre attractiveness', *International Journal of Retail and Distribution Management*, Vol.29 No.2, pp.76-86.

Zeithaml, V.A., Berry, L.L. and Parasuraman, A. (1996), 'The Behavioral Consequences of Service Quality', *Journal of Marketing*, Vol. 60 (April), pp. 31-46.

Table I: Demographic profile

Demographic Factors	Valid Items	Frequency	Percentage

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Gender	Male	267	44.4
	Female	327	54.3
Age (in Years)	18-27	163	27.1
	28-37	159	26.4
	38-47	102	16.9
	48-57	68	11.2
	58-67	64	10.6
	68 and above	37	6.1

Table II: Cronbach alpha values

Constructs	Cronbach's Alpha
Healthscape	0.875
Personnel	0.869
Hospital Image	0.739
Trustworthiness	0.805
Clinical Care Process	0.823
Communication	0.869
Relationship	0.845
Personalization	0.730
Administrative Procedures	0.896
Perceived Service Quality	0.873
Customer Satisfaction	0.811
Behavioural Intentions	0.882

Table III: Bivariate correlations among the dimensions

	HS	PL	HI	TW	CCP	COM	REL	PER	AP	PSQ	SAT	BI
HS	1											
PL	.668	1										
HI	.509	.574	1									
TW	.603	.702	.508	1								
CCP	.600	.675	.604	.631	1							
COM	.605	.673	.505	.646	.730	1						
REL	.420	.462	.386	.454	.472	.567	1					
PER	.446	.494	.423	.418	.523	.556	.582	1				
AP	.526	.565	.450	.586	.581	.693	.520	.521	1			
PSQ	.55	.551	.483	.548	.564	.587	.445	.414	.531	1		
SAT	.532	.604	.546	.585	.617	.591	.412	.451	.528	.633	1	

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

BI	.518	.575	.534	.527	.587	.562	.384	.441	.462	.651	.765	1

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

HS - Healthscape, PL - Personnel, HI - Hospital Image, TW- Trustworthiness, CCP - Clinical Care Process, COM - Communication, REL - Relationship, PER – Personalization, AP - Administrative Procedures, PSQ – Perceived Service Quality, SAT – Satisfaction,
BI – Behavioural Intentions.
All are positive and significant at 0.001.

Table IV: One-way ANOVA for assessing convergent validity – independent variable service quality

Descriptive Variables	Degrees of freedom		F	Sig.
	Between Groups	Within Groups		
Healthscape	13	588	24.219	.001
Personnel	13	588	28.381	.001
Image	13	588	23.463	.001
Trustworthiness	13	588	28.694	.001
Clinical Care Process	13	588	29.622	.001
Communication	13	588	31.435	.001
Relationship	13	588	17.029	.001
Personalization	13	588	13.205	.001
Administrative Procedures	13	588	25.967	.001

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Table V: Validity tests

Validity	Measures used	Correlation coefficient (r)
Convergent validity	Perceived service quality	0.665 (p<0.001)
Predictive validity	Repurchase intentions	0.621 (p<0.001)

Table VI: Collinearity statistics

Independent Variable	Tolerance*	VIF*
Healthscape (HS)	.475	2.104
Personnel (PL)	.347	2.879
Hospital Image (HI)	.567	1.765
Trustworthiness (TW)	.413	2.421
Clinical Care Process (CCP)	.353	2.831
Communication (COM)	.308	3.248
Relationship (REL)	.559	1.787
Personalization (PER)	.544	1.838
Administrative Procedures (AP)	.455	2.198

*assessed using Enter regression method

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Table VII: Model summary

R	R Square	Adjusted R Square	F	Sig.	Durbin-Watson
.677	.458	.451	71.65	.001	1.987

Table VIII: Split data model summary - estimation sample – 0

R		R Square	Adjusted R Square	F	Sig.	Durbin-Watson	
Split = 0.00 (Selected)	split ~= 0.00 (Unselected)					Split = 0.00 (Selected)	split ~= 0.00 (Unselected)
.696	.623	.484	.478	75.059	.001	2.000	2.053

Table IX: Split data model summary - validation sample – 1

R		R Square	Adjusted R Square	F	Sig.	Durbin-Watson	
Split = 1.00 (Selected)	split = 1.00 (Unselected)					Split = 1.00 (Selected)	split ~= 1.00 (Unselected)
.664	.652	.441	.433	53.618	.001	2.056	1.944

Table X: Summary - Validation results: stepwise regression

	Entire data	Estimation sample	Validation sample
ANOVA significance (sig<=0.05)	<0.001	<0.001	<0.001
R^2	0.458	0.484	0.441

Table XI: Reliability and validity tests

Reliability	Content Validity	Criterion related Validity	Construct Validity
split-half reliability tests	Literature	Predictive validity (2 tests)	Convergent validity (2 tests)
inter-item consistency reliability	Pilot testing	Concurrent validity (2 tests)	

Table XII: Reliability and validity a comparison with established scales

Sl. No.	Particulars	Parasuraman *et al.,* (1988)	Aagja and Garg (2010)	Duggirala *et al.,* (2008)	Pai and Chary (2016)

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

1	Instrument	SERVQUAL	PubHosQual		
2	Dimensional Structure	Five	Five	Seven	Nine
3	Application Domain	All services	Public Hospitals	Hospital Services	Hospital Services
4	Reliability	$0.72 < \alpha < 0.86$	$0.58 < \alpha < 0.89$	$0.775 < \alpha < 0.906$	$0.73 < \alpha < 0.896$
5	Content Validity	Measured	Measured	Measured	Measured
6	Construct Validity	Convergent Validity	Convergent and nomological validity are successful; discriminant validity had mixed results.	Convergent Validity	Convergent Validity
7	Criterion related validity			Concurrent validity	Both concurrent and predictive validity are measured

Appendix 1: Questionnaire

1. Modern and up-to-date equipment (e.g., Computerized tomography (CT) and Magnetic Resonance Imaging (MRI), patient information and billing) to serve patients more effectively.
2. Physical facilities are visually appealing.
3. Facility adequacy (e.g., wards, beds, operation theatre, Intensive Care Unit (ICU)).
4. Cleanliness (toilets, rooms and wards).
5. Infection-free environment/treatment.
6. Hospital staff follow hygienic care and procedures (e.g., wearing gloves).
7. Employees are dressed neatly.
8. Drugs are available.
9. Comfortable ambient conditions with proper lighting.
10. The hospital has an appealing atmosphere.
11. The hospital has clean rooms without foul smell.
12. There are sufficient waiting areas for patients and families.
13. It is easy to find my way in the hospital.
14. It is easy to find care facilities (laboratory, Doctor's office).
15. It is easy to use the amenities (public telephone, cafeteria, etc.).
16. Courtesy is shown by hospital staff towards patient and patient party.
17. Doctors and nurses are available as and when required.
18. The doctors are competent and skilful.
19. Knowledgeable nurses.
20. Paramedical and support staff are competent.
21. The doctors are friendly, caring and understand patient's feelings and needs.
22. Doctors talked to me frankly and politely.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

23. Nurses give prompt and timely attention.
24. Nursing staff are polite and well-mannered.
25. Staff at the hospital are polite.
26. Doctors are professional.
27. Good doctors are available in the hospital.
28. The hospital has positive reputation.
29. Sincerity, honesty and ethics followed in providing medical services.
30. The hospital staff run various programs for patients to support different societal sections.
31. The hospital provides medical services with nominal cost to the needy patients.
32. Patient privacy and confidentiality are maintained by hospital staff.
33. Hospital staff provide services as promised and on time.
34. Equal treatment for all.
35. Confidence in the doctor who treated you in the hospital.
36. The hospital provides patients with services beyond medical treatment.
37. Faultless assessment of health conditions by doctors.
38. Explanations are provided by the doctor about health status, medical tests.
39. Treatment procedures and outcomes are explained.
40. Medical advice and instructions provided by doctor.
41. Diagnosis is only made after careful examination.
42. Doctors spent enough time examining the patient.
43. Doctors provide information quickly.
44. My family was told what they needed to know.
45. Hospital staff provided adequate information about my illness/treatment(s).
46. Information can be easily obtained.
47. Obtaining information from hospital administrative personnel (e.g., admission, treatment, discharge) is easy.
48. Extent to which doctors' answer patient's questions and explain treatment that I could understand.
49. I feel good about the interaction I have with doctors at the hospital.
50. I feel good about the interaction I have with nurses at the hospital.
51. I feel good about the interaction I have with staff at the hospital.
52. I have built a close relationship with some staff at the hospital.
53. I have built a close relationship with the doctor at the hospital.
54. I have built a close relationship with nurses at the hospital.
55. I always get personalized attention from staff at the hospital.
56. Hospital staff treat me as a human being and not just a patient.
57. The doctor calls my name while addressing me.
58. Waiting time to consult the doctors is minimum.
59. Staff provide right patient services the first time, every time.
60. Reasonable waiting time spent for diagnostic test and treatments.
61. Time between the two successive processes is minimum.
62. The process for setting up the appointment was simple and easy.
63. Appointments at the hospital run on time.
64. The hospital's records and documentation are error free.
65. The interaction among department staff is well managed.
66. I believe the hospital is well-managed.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Perceived Service Quality (Brady and Cronin, 2001; Dagger et al., 2007; Parasuraman et al., 1988)
67. The overall service quality provided by hospital staff is excellent.
68. The overall service quality provided by hospital staff is impressive.
69. The service provided by the hospital is a high standard.
70. I believe the hospital offers service that is superior in every way.
Satisfaction (Dagger et al., 2007; Greenfield and Attkisson 1989; Oliver 1997)
71. My feelings towards the hospital are very positive.
72. I feel good about coming to this hospital for my treatment.
73. Overall, I am satisfied with the hospital and the service it provides.
74. I feel satisfied that the treatment results are the best that can be achieved.
75. The extent to which my treatment has produced the best possible outcome is satisfying.
Behavioral Intention (Dagger et al., 2007; Headley and Miller 1993; Taylor and Baker 1994; Zeithaml et al., 1996)
76. If I had to start treatment again, then I would want to come to this hospital.
77. I would highly recommend the hospital to other patients.
78. I have said positive things about the hospital to my family and friends.
79. I intend to continue having treatment, or any follow-up care I need, at this hospital.
80. I have no desire to change hospital.
81. I intend to follow the medical advice given to me at the hospital.
82. I am glad to have my treatment at this hospital rather than somewhere else.

Downloaded by INSEAD At 20:59 24 January 2018 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.