

Validation and adjustment of the Patient Experience Questionnaire (PEQ) based on a Norwegian hospital study.

SETH AYISI JUNIOR ADDO (✉ seth.a.addo@uis.no)

University of Stavanger

ESPEN OLSEN

University of Stavanger

REIDAR JOHAN MYKLETUN

University of Stavanger

Research Article

Keywords: patient experiences, PREMs, psychometrics, CFA, hospital, service climate, second-order factor, Norway

Posted Date: March 10th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-223253/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background- This paper assesses the psychometric qualities of the Patient Experience Questionnaire (PEQ), thereby validating a patient-oriented service climate measurement model in a hospital environment, and modifies the model based on empirical results. This study employed survey data gathered by the Norwegian Institute of Public Health from adult inpatients at somatic hospitals in the Health South-East RHF in Norway. The survey engaged 4,603 patients out of 8,381 from 5 main hospitals in the region. Data was analysed with the aid of AMOS, SPSS, and Microsoft Excel.

Results- The study found that an 8-factor model of the PEQ generally showed good fitness to the data, but assessment of discriminant validity showed that this was not the optimal factor solution among 4 of the 8 dimensions. After comparisons of models, support was finally found for a model with a second-order factor for 4 of the factors: *"nurse services"*, *"doctor services"*, *"information"*, and *"organisation"*, collectively named *"treatment services"*. The proposed model demonstrated good validity and reliability results.

Conclusions- The results present theoretical and practical implications. The study recommends that inferential analyses on the PEQ should be done with the second-order factor. Also, a revision of the PEQ is recommended subject to more confirmatory studies with larger samples in different regions. The study indicates a second-order factor structure for assessing and understanding patient experiences, a finding having both theoretical and managerial implications.

Introduction And Background

Healthcare professionals are facing heavy pressure to meet the growing needs and expectations of patients (1). This owes to the increasing and alarming rate of morbidity and multi-morbidity in Western countries (2), together with aging populations and the healthcare needs of the aged. This pressure on healthcare professionals has increased in recent times with the outbreak of global pandemics such as COVID-19. Notwithstanding these morbidity rates and growing needs of patients, healthcare providers and professionals are expected to ensure positive patient experiences. This study, focusing on hospitals and their professionals, seeks to examine the patients' experiences with hospital service climate, focusing on the psychometric quality of a patient-reported experience measure (PREM).

The endeavour of gathering patients' experiences with healthcare has gained popularity, thus resulting in the development of PREMs that have been used in surveys in various countries. In a bid to clarify the meaning of patient experiences, Wolf and Jason (3) synthesised various definitions of the concept and maintained that patient experiences comprise individual as well as collective events and occurrences that happen in the process of caregiving, and this has strong links with the patients' expectations and how they were met. Wagland, Recio-Saucedo (4) noted that significant progress has been made in understanding patient experience. The concept is viewed as interactions of patients with aspects of the healthcare delivery such as nurse services, doctor services, organisation of the caregiving process in

hospitals, and information delivery, where these aspects (dimensions) culminate in the entire continuum of experience that patients have with healthcare, as reported by the patients.

From patients' perspective, interactions with dimensions of healthcare have been theoretically underpinned by the Donabedian (5) framework for assessing healthcare quality (6) and considered the most widely used in the healthcare sector to assess quality (7). According to this framework, quality of healthcare can be assessed by making inferences under 3 categories: structure, process, and outcome. The structure deals with the setting in which care is given, for instance, facilities, equipment, and human resource. The process deals with what is done in giving and receiving care, for instance, nurse and doctor services as well as good communication and information sharing between patients and hospitals; lastly, the outcome deals with the effects of care on health and well-being (5, 7).

Increased understanding of patient experiences of hospital climate has similarly been aided by increased research and several studies on measuring the construct. Measurements in social science provide adequate guidelines for assessing phenomena and people's attributes that are not directly and easily observable (8). Employing poor and inadequate measures can be very costly to research, in terms of drawing invalid conclusions, making policy decisions based on false information, and wasting respondents' time and efforts (9). DeVellis (9), however, indicated that a major challenge to developing adequate measures in social science is the immaterial nature of social science constructs supported by constantly changing theories. This makes measurements in social science susceptible to constant changes in performance and adequacy in assessing the constructs. Consequently, social science measures need to be constantly reviewed and reassessed to keep them abreast with changing theories and constructs and to uphold their validity and reliability. Therefore, reassessing PREMs to ensure adequate psychometric qualities is essential for theoretical and practical advancement of knowledge of patients' experiences, hence the focus and aim of this study.

Justification of the study

The goal to accurately measure patient experiences has resulted in several PREMs for general and specialised healthcare. The questions and dimensions that these PREMs have produced are indicative of patients' shared experiences. Most of these measures identified similar dimensions of experiences, such as those relating to nurse services, doctor services, information and communication, hospital organisation and standards, and discharge from the hospital (10-14). Although some of these studies differed with regard to the naming of the dimensions, the content of the items remained very similar among the PREMs. This study is underpinned by 3 main justifications: (i) psychometric statistical analyses have evolved over the years with more robust tools in validating scales; (ii) due to the plethora of patient experience measures and unascertained psychometric qualities, existing PREMs should be re-examined to ascertain their validity and reliability, rather than developing new ones; (iii) there is the need for scrutiny of existing PREMs because they may be prone to theoretical and practical changes. These justifications are elaborated below.

The Norwegian Institute of Public Health (NIPH) conducted a survey in the East health region among a few hospitals, adapting an earlier validated PREM, the Patient Experience Questionnaire (PEQ) (15). In the development and validation study, Pettersen, Veenstra (15) employed literature reviews, focus groups, pilot studies, and 2 cross-sectional surveys (1996 and 1998) across 14 hospitals in Norway. The study used exploratory factor analysis, reliability test (Cronbach's alpha), and construct validity test. The study found 10 factors and 20 final items out of an initial 35 items: *"information on future complaints"*, *"nursing services"*, *"communication"*, *"information examinations"*, *"contact with next-of-kin"*, *"doctor services"*, *"hospital and equipment"*, *"information medication"*, *"organisation"*, and *"general satisfaction"*. All the factors recorded Cronbach's alpha scores between 0.61 and 0.83. Construct validity was also ascertained in the study by examining the relationship between the instrument and demographic factors such as age and gender. Stressing the lack of valid and reliable instruments, Pettersen, Veenstra (15) concluded that it is imperative to re-examine existing patient experience measures so as to improve methodology. They further recommended employment of the PEQ for future in-patient experience surveys, hence the choice for the current study. Although this measure was adapted and modified for use by the NIPH, the performance of the measure should be called into question because this measure was developed and validated more than a decade ago. Psychometric analyses are evolving with more robust validating tools and methods, and this is evident in the study by Pettersen, Veenstra (15) where issues such as discriminant validity and measurement invariance as well as other psychometric issues were absent in the analyses—a gap that the current study tackles.

Beattie, Murphy (16) also noted the problem of multiple patient experience measures with unascertained psychometric quality. This problem has hindered the use of data from patient experience surveys to adequately improve and sustain quality of care in hospitals. In the systematic review, Beattie, Murphy (16) developed a matrix to help choose PREMs for research and to identify research gaps in existing ones. This matrix showed that the PEQ study by Pettersen, Veenstra (15) lacked analyses such as criterion-related validity. On this basis, the current study asserts that rather than developing more PREMs (which seem already saturated), existing ones should be re-examined, as recommended earlier by Pettersen, Veenstra (15), in light of current analyses and conceptual underpinnings. This need for re-examination has also been recommended by other systematic reviews on patient experience (17, 18).

Additionally, some PREMs have been developed in Norway to capture the phenomenon of patient experiences with general health practice as well as experiences with specific health issues and fields, with most of them asking questions on general patient satisfaction (e.g. 12, 15, 19, 20). DeVellis (9) noted that the constantly changing theories of social science constructs challenge the adequacy of their measures. Haugum, Danielsen (17) similarly recommended the need to repeat patient experience surveys and their outcomes in order to generate more validated instruments, as they are potentially affected by constantly changing contextual factors. By inference, it can be said that the underlying psychometric rigors of a PREM can dwindle as they are employed over a long period. Although several surveys exist on patient experiences on various issues (e.g. 21, 22-25), a re-analysis of the psychometric performance of any particular measure is lacking. The quest to improve healthcare delivery and hospital service climate based on patients' experiences should begin with ascertaining the psychometric quality of PREMs. Based

on these justifications, the purpose of this article is to test the psychometric qualities of the PEQ, thereby validating a patient-oriented service climate measurement model in a hospital environment.

Conceptualisation of patient-oriented hospital service climate

The concept of 'climate' has been used extensively in the organisational setting for decades as an indicator of conditions and shared perceptions from organisational members, mainly employees. The concept is defined as comprising measurable aspects of the work environment that are perceived and shared as the formal and informal practices, policies, and procedures by individuals in an organisation (26, 27). Thus, these perceptions are socially constructed and are shared among the individuals (28); furthermore, they have an influence on individuals' motivations and actions (26). A number of sub-concepts of organisational climate have been developed, including safety climate (29) and service climate (30). Service climate is explained as the shared perceptions of employees regarding service quality based on policies, practices, and procedures (31), reflected in feedback from customers on service and satisfaction, among other elements (30). These concepts are applicable and measurable in different industries and sectors. In the health sector, organisational climate and sub-concepts have been explored and examined mainly from the workers' perspective to indicate employee conditions, quality of patient care, and patient safety (e.g. 32).

This study links organisational climate with Donabedian's framework to conceptualise hospital service climate, not from the employees' perspective but from the patients' perspective. Thus, in this study, hospital service climate is defined as patients' shared experiences and perceptions of hospital work environments, the services, and the formal and informal practices and procedures that make up the entire caregiving process, which inform patient outcomes such as satisfaction, health benefits, and health level. In sectors such as health, the perceptions of patients are imperative for managerial and organisational development. Patients' reports on their experiences with the hospital climates, such as the structures and processes of giving care, can therefore be useful indicators and predictors of their outcomes, evidenced in previous studies (e.g. 33, 34, 35). This affords hospital management adequate information on their service climate in order to improve the quality of healthcare. The study pursues its aim in light of this theoretical linkage, based on the idea that adequate PREMs generate adequate information on patients' experiences with hospital climates and resulting outcomes.

Methods

Sample and data collection

This study employed anonymous secondary survey data from the Norwegian Institute of Public Health gathered from adult inpatients at somatic hospitals in the Health South-East RHF in Norway. The survey engaged patients from 5 main hospitals in the region and who were admitted for at least a day. As part of eligibility criteria, patients who were admitted between October and November in 2015 were included. The survey was started by the Norwegian Knowledge Centre for Health Services in the fall of 2015 and was continued at the Norwegian Institute of Public Health in the first quarter of 2016. Patients who visited the

5 hospitals were identified through their contact information after they were discharged. Questionnaires were sent to their respective addresses via post mail with a return envelope. About 8,381 patients were eligible and contacted. The total number of respondents who completed and returned their questionnaires was 4,603, yielding a response rate of 54.92%. Patients were asked to consider various aspects of their experience being admitted. The questionnaire aimed at using feedback to identify which areas are working well and which areas the hospital should work to improve.

Instrument

The Patient Experience Questionnaire (PEQ) comprised 8 dimensions and 33 items (see Table 5) as well as items on patient safety, patient satisfaction, and overall health benefits and health level. The NIPH adapted the questions for the survey from the PEQ developed and validated by Pettersen, Veenstra (15): *“nurse services”* (items N1-N7), *“doctor services”* (items D1-D7), *“information”* (items IF1-IF3), *“organisation”* (items OR1-OR4), *“next of kin”* (items NK1 and NK2), *“standard”* (items S1-S6), *“discharge”* (items DC1 and DC2), and *“interaction”* (items IT1 and IT2). These items were measured on a 5-point Likert scale ranging from ‘Not at all’ (1) to ‘To a very large extent’ (5). Patient safety was measured with 12 items, while patient satisfaction, health benefit, and health level were measured with 1 item each. Background information, such as questions on whether or not the patient chose the hospital he/she was admitted to, was also included in the questionnaire.

Data analysis

Preliminary analyses

The study analysed the data with the aid of Microsoft Excel, SPSS v.24, and AMOS v.25. Preliminary analysis (such as checking for normality, outliers, and missing value analysis) was conducted in SPSS. In order to ensure maximum privacy of respondents and still maintain relevant variables for analysis, departments for the analysis were aggregated into medical departments (*Med*) and surgical departments (*Kir*) across the hospitals based on the more specific and varied information on units in the hospitals provided by participants. This aggregation was performed according to the departmental codes for health institutions provided by the Norwegian Health Authority.

Measurement model development

The initial measurement model (see Model 1 in Table 3) was developed in AMOS without modification indices (due to the exclusion of missing values). Missing values were replaced after the estimation of the initial model to obtain modification indices for correlating error terms among the items and improving the fitness of the model (see Model 2 in Table 3). Due to the non-randomness of the missing values, being mindful of how they were replaced was necessary. The study chose to use multiple imputations to replace them as recommended for non-randomness (36, 37). However, the 5 different imputations generated could not be pooled in AMOS as a single imputation for the estimation of the model. Thus, the missing values were eventually replaced with the series mean method. It is noteworthy that the missing

values were only replaced in order to generate a full estimation with modification indices for correlating the error terms. Although all subsequent models after the initial model were estimated with the correlated error terms, estimations were done on the data with missing values, with the aim of obtaining a more accurate fit of the data to the models.

The initial model with modifications (correlated error terms), Model 2, was compared to 6 other models (see Models 3-8 in Table 3), obtained by combining some dimensions into a single factor to further justify the fitness of the modified initial model. These combinations were based on the correlation coefficients between the dimensions (see Table 3 for the combinations of the dimensions). In addition, a proposed model containing a second-order factor for *“nurse services”*, *“doctor services”*, *“information”*, and *“organisation”* was also developed and compared to the initial modified model based on the validity tests, correlation analyses, and theoretical justifications (wording of questions). Fitness of all the models was ascertained using the following indices: Comparative Fit Index (CFI), Tucker Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and the PCLOSE. The thresholds recommended by Hu and Bentler (34) are presented in Table 1.

Validity and reliability

Validity in this study was ascertained using convergent, discriminant, and predictive validity tests. Convergent validity deals with the relationship between a latent construct (patient experience dimensions) and its items (38). The Average Variance Extracted (AVE) was used to check convergent validity, where values must be at least 0.50, indicating that at least half of the variance in the construct (dimension) is explained by its items. Discriminant validity focuses on a construct and its items in relation to other constructs—that is, how different one construct (or dimension) and its items are from other constructs in the model (38). Discriminant validity was examined using the Fornell-Larcker procedure, where discriminant validity is supported when the square root of the AVEs is greater than the correlation coefficients between the constructs (39). Predictive validity focuses on the ability of the measure and dimensions to relate to and predict previously ascertained outcomes in literature. This was determined through correlation and regression analyses between patient experiences (and dimensions) and outcome variables (patient satisfaction, health benefits, and health level) with the aid of SPSS. Reliability of the measurement model was also determined using composite reliability values for every dimension of the patient experience measure, with a recommended value of at least 0.70.

Ethical considerations

This study, with regard to data collection, analysis, and compilation, was conducted within the ethical and legal provisions and guidelines of the Norwegian Institute for Public Health (NIPH) and the University of Stavanger. The Norwegian Data Protection Authority and the Norwegian Directorate of Health approved the procedures in the survey. The hospital data protection official assessed the data processing in the hospitals where survey extension took place. Informed consent was obtained from participants in the survey. Respondents were informed that participation was voluntary and they were assured of confidentiality of the information they will provide. Respondents were also informed that they could opt

out of the survey at any point as well as the procedure for opting out if they wished. Data was stored in a safe repository with a password, only accessed by the researchers. This study did not present results that revealed patients' identities, thus maintaining anonymity of respondents and confidentiality of responses. All relevant ethical requirements were duly upheld.

Results

Preliminary analysis and sample characteristics

The study made use of responses from 4,603 participants. Outliers were recorded for some of the questions, but this was to be expected considering the varied background characteristics, such as age and number of days spent in the hospital, which could influence participants' experiences. Nonetheless, most of these outliers were not deemed extreme based on the 1.5 and 3.0 interquartile ranges. Normality was also ascertained, using the -2 and +2 range (40), for all items of patient experience, except the kurtosis value for 1 item on *"nurse services"* and 1 item on *"doctor services"*. Overall, the data could be said to be normally distributed to a large extent. The sample for the study was taken from 5 hospitals and characterised by a somewhat fair age distribution of patients across 3 groups: 60 years and below, between 61 and 73 years, and 74 years and above. Most of the respondents were admitted for 3 or fewer days, and more of them were also admitted to the medical department aggregate (Med). Table 2 presents the sample characteristics for this study.

Initial measurement model development, modifications, and comparisons

The initial CFA model (Model 1 in Table 3), with the 8 dimensions of patient experience, was then developed to be tested. The model showed acceptable fitness to the data based on fitness indices. Nonetheless, there was a need to improve the fitness through modifications in order to reduce measurement errors and to obtain more accurate loadings of the observed items on their dimensions. Some modifications were made by drawing covariance between some error terms on the same dimensions with the rationale that, by virtue of sharing commonalities on the dimension, they are more justified to share similar error terms, thus reducing duplications of random measurement error of items. In total, 19 modifications were made based on the covariance coefficients, with the highest coefficient as 895.667 between S2 and S4 (*"standard"*) and the lowest as 40.390 between D4 and D7 (*"doctor services"*). Aside from the coefficients, these modifications were theoretically justified. For example, the item D2 was worded, *"Did you find that the doctors took care of you?"*, and D4 was worded as *"Did the doctors have time for you when you needed it?"* Participants may have given closely related responses due to the phrases *"taking care"* and *"having time when you needed"*; therefore, it was no surprise that they shared similar error terms, leading to considerable covariance coefficient. These statistical and theoretical justifications were made for each covariance drawn. The most modifications were made to *"doctor services"* (seven), followed by *"standard"* (five), *"nurse services"* (four), *"information"* (two), and *"organisation"* (one). No modifications were made to *"next of kin"*, *"discharge"*, or *"interaction"*, owing to very low covariance coefficients (below 20). The initial model with these modifications (Model 2) thus

produced excellent fitness values for all indices. Furthermore, the model was compared to 6 other models (explained under Methods section), where the initial model with modifications showed the best fitness to the data. The fitness indices of the initial model before and after modifications, as well as those of the 6 alternative models for comparisons, are presented in Table 3.

Measurement invariance across hospital departments aggregated into two groups

Model 2 was further examined for invariance across 3 categories: configural, metric, and scalar. Measurement invariance tests seek to ascertain whether the measurement model differs across variant groups in a data. The goal is to achieve little or insignificant variance across these groups in order to inspire confidence in the ability of the measure to generate accurate responses and assessments across groups (41). Configural invariance results (see Model 9 in Table 3) showed that the model had acceptable-to-excellent fitness to the data, thus ascertaining configural invariance for the 8-factor patient experience measure across the 2 hospital department aggregates. With regard to metric invariance, the chi-square test showed that the fully constrained model and the unconstrained model were different across the department groups and, thus, not metrically invariant. However, MacKenzie, Podsakoff (42) maintained that “full metric invariance is not necessary for further tests of invariance and substantive analyses to be meaningful, provided that at least one item (other than the one fixed at unity to define the scale of each latent construct) is metrically invariant” (p. 325). Thus, the critical ratios test was performed to examine whether the dimensions and the items were metrically invariant enough for further meaningful analyses. The analysis revealed that for all dimensions, with the exception of “*next of kin*”, there was at least 1 item that was not statistically significant (metrically invariant) besides the item that was constrained for that dimension in the model. This means that the 2 items on the “*next of kin*” dimension had significantly different loadings (parameters) across the aggregated departments. Nonetheless, this test showed the model was metrically invariant across the departments to a large extent. The results of this test are presented as a supplementary table. Scalar invariance was then examined for the model based on the differences in the measurement intercepts. The analyses showed that the model did not have scalar invariance. Differences in intercept estimates of items between the departments were computed, showing that almost all the items did not have scalar invariance across the 2 departments. The results are presented as a supplementary table.

Reliability

Reliability for the measure was ascertained using composite reliability (CR) values. Generally, CR values above 0.70 are deemed acceptable to justify reliability. From Table 4, it is seen that all the dimensions recorded CR values above 0.70, with the highest being “*doctor services*” (0.92) and the lowest being “*interaction*” (0.72).

Convergent and discriminant validity

Convergent validity was examined using the AVE values, where an AVE value of at least 0.50 is considered acceptable (38). Table 4 shows that all dimensions, with the exception of “*standard*”, recorded

values above 0.50, thus ascertaining convergent validity. Discriminant validity was ascertained using the Fornell-Larcker procedure. There, discriminant validity is supported when the square root of the AVEs is greater than the correlation coefficients between the constructs (39). From Table 4, it is seen that discriminant validity issues were observed for *“doctor services”* (in relation to *“information”*); *“organisation”* (in relation to *“doctor services”*, *“nurse services”*, and *“information”*); and *“standard”* (in relation to *“organisation”*). This means that these 3 dimensions were not distinct from the others enough for each to measure the different sub-concepts under patient experience.

Construct validity, item loadings, and deletion

Construct validity for the items was examined by checking item loadings (parameter estimates) on their dimensions. Generally, good loadings were recorded as a majority of the items had loadings above 0.60. The item loadings ranged from 0.88 (on *“discharge”*) to 0.55 (on *“standard”*). Two items had loadings below 0.60: 0.58 (ORG 2) and 0.55 (ST 5). Based on the suggestion of the master validity tool (43), these items together with a third (ST4) were deleted in a bid to boost the validity of the measure. Item loadings before and after deletion are presented in Table 5. After deletion, the dimension *“standard”* recorded an increase in AVE value, indicating that the remaining 4 items explained more variance in the dimension than the original 6 items, seen in Table 6. Figure I presents the model after item deletion as well as validity and reliability checks. See Model 10 in Table 3 for the fit indices of this model.

Criterion-related validity

The study then assessed the predictive validity of the model based on its ability to relate to and predict outcome variables ascertained in existing literature. Overall satisfaction, health benefits, and health level were used as outcome variables while the patient experience measure and its dimensions were used as predicting variables. Patient experience measure and dimensions were computed with retained items after item deletion, and standard multiple regression was performed. The results showed that overall patient experience and each individual dimension related to and predicted at least 1 outcome variable positively and significantly. These results are presented in Table 7.

Proposed measurement model

A proposed model (see Model 11 in Table 3) was developed, taking into consideration the frequencies of missing values for the items and the discriminant validity concerns. Items with missing values of more than 20% were excluded; therefore, the dimensions of *“discharge”* and *“interaction”* were removed from the model. The items on *“next of kin”* had more than 20% but the dimension was maintained. The questions were the following: *“NK1: Were your relatives well received by the hospital staff?”* and *“NK2: Was it easy for your relatives to get information about you while you were in the hospital?”* These questions were maintained because, unlike the other dimensions, relating and answering them depended on factors that are largely beyond the control of the patient, such as whether or not the patient had any relatives alive who visited the hospital and whether the patient stayed in the hospital long enough for relatives to visit the hospital. A second-order factor was added in the proposed model for *“nurse services”*,

“doctor services”, “information”, and “organisation”, collectively labelled “treatment services”. This was based on the discriminant validity results, correlations among them, and the nature of the questions asked under these dimensions. The 2 lowest loading items (ORG 2 and ST 5) that were previously deleted were still excluded from this model. The proposed model showed excellent fitness to the data (similar to Model 10) and also met convergent, discriminant, and criterion-related validity requirements. See Figure II for the proposed model. Table 8 presents comparisons of tools and findings between the validation study by Pettersen, Veenstra (15) and the current study.

Discussion

This study presents some major findings. First, the study confirmed that the 8-factor model showed good fitness to the data. The model achieved configural and metric invariance but not scalar invariance. The study also found that reliability values were all acceptable and all the dimensions, except *“standard”*, attained the recommended 0.50 AVE value for convergent validity. With regard to discriminant validity, *“doctor services”* (in relation to *“information”*), *“organisation”* (in relation to *“doctor services”, “nurse services”, and “information”*) and *“standard”* (in relation to *“organisation”*) had issues. Construct validity and criterion-related validity were supported for majority of the results. Finally, a model including a second-order factor was proposed. The second-order factor, named *“treatment services”*, consisted of 4 first-order factors: *“nurse services”, “doctor services”, “information”, and “organisation”*. Moreover, the dimensions of *“standard”* and *“next of kin”* were included in this final model, but *“discharge”* and *“interaction”* were excluded. Hence, the final model included one second-order factor comprising 4 sub-factors as well as *“standard”* and *“next of kin”*.

The dimensions with associated items found in this study were similar to those found by Pettersen, Veenstra (15) while some dimensions, such as *“doctor services”, “nurse services”, “organisation”, “information”, and “hospital standards”,* overlapped with dimensions found by other studies (12, 13, 20). Invariance tests conducted in the present study were absent in the study by Pettersen, Veenstra (15), which marks a good contribution of this study. The tests showed that the model achieved invariance across the aggregated departments with regard to structure and pattern (configural) as well as the loadings of the items on their respective dimensions (metric). However, scalar invariance was not achieved for this model. Considering the diverse nature of the sample, as well as the aggregation of the departments into broad categories, this finding was expected. Putnick and Bornstein (44) asserted that scalar invariance is the most stringent compared to configural and metric, and instances of rigid scalar non-invariance could mean that the construct is generally variant across different groups. The findings also showed that reliability was good, based on composite reliability values, similar to the Cronbach’s alpha values obtained by Pettersen, Veenstra (15).

With regard to validity tests, the study found that all the dimensions, except *“standard”*, attained the recommended 0.50 AVE value for convergent validity, similar to other related studies that examined similar dimensions using other instruments (45). However, discriminant validity issues were found for *“doctor services”* (in relation to *“information”*), *“organisation”* (in relation to *“doctor services”, “nurse*

services", and *"information"*) and *"standard"* (in relation to *"organisation"*). Discriminant validity was also missing in the study by Pettersen, Veenstra (15), thus indicating another good contribution of this study. Examining the wordings of their items gives some possible explanation for this finding. For instance, D1 under *"doctor services"* was worded as *"Did the doctors talk to you so you understood them?"*, while questions under *"information"* included *"IF2. Did you know what you thought was necessary about the results of tests and examinations?"* and *"IF3. Did you receive sufficient information about your diagnosis or your complaints?"* It is highly likely that patients will receive information on results and diagnosis mainly from their doctors and, as such, answering questions under *"information"* may be significantly influenced by the perception of how well the doctors spoke to these patients. Similarly, questions under *"organisation"* were *"OR1. Did you find that there was a permanent group of nursing staff that took care of you?"*, *"OR2. Did you find that one doctor had the main responsibility for you?"*, *"OR3. Did you find that the hospital's work was well organized?"*, and *"OR4. Did you find that important information about you had come to the right person?"* These questions feature clear wording relating to *"nurse services"*, *"doctor services"*, *"information"*, and *"standard"*, and it is therefore not surprising that no clear distinctions were found among them as constructs. Construct validity was also achieved with a majority of the items recording loadings of above 0.60. This was also achieved in the validation study by Pettersen, Veenstra (15) using a different method and in related studies using other instruments with similar dimensions (12, 13). One item on *"standard"* and one on *"organisation"* were, however, deleted due to loadings below 0.60, while another on *"standard"* was deleted in a bid to improve the discriminant validity. Perhaps the wording of these questions made them difficult for patients to understand clearly and respond accordingly. For instance, item S5 was framed as *"Was the food satisfactory?"* Patients may be left to decide what is meant by 'satisfactory', thus making the question too vague, or perhaps the different dietary requirements and preferences made this question more loosely defined. Again, item OR2 was framed as *"Did you find that one doctor had the main responsibility for you?"*, a question probably dependent on the ailments of the patient and likely to be out of the control of hospital organisation. Thus, if a patient's ailments require more than a single main doctor, then this question may suggest to the patient that having 2 or more main doctors reduces the ability of the hospitals to organise their work well. Criterion-related validity was ascertained for the overall measure as well as the dimensions in predicting at least 1 of the 3 outcome variables: satisfaction, health benefits, and health level, which is consistent with previous studies (e.g. 33, 34, 35).

Lastly, a model with a second-order factor, *"treatment services"*, for 4 of the dimensions was proposed based on the results of the validity and reliability analyses: *"nurse services"*, *"doctor services"*, *"information"*, and *"organisation"*. This constitutes the most important contribution of this study since this possibility was not explored in the study by Pettersen, Veenstra (15), perhaps owing to the absence of discriminant validity examinations in their study, and since this indicates a change in the factor structure of the PEQ. Rindskopf and Rose (46) observed that second-order factors reflect relationships among first-order factors. It is worth noting that related studies that developed other PREMs for generic and specific health issues also found these 4 dimensions in common (e.g. 11, 13, 20, 47). Although these studies did not develop a second-order factor for these dimensions, this is indicative of the prominence of these 4

variables in measuring and understanding patient experiences. The current finding, therefore, builds on this prominence to illustrate the high interrelationships and inextricable links among these factors, which brings some theoretical and practical implications to the fore.

Theoretical Implications

This study brings a very important, yet mostly ignored, contribution to the patient experience and quality healthcare literature: a need for more validation studies and surveys on patient experiences. The study responds to the recommendation by Pettersen, Veenstra (15) that existing PREMs need scrutiny and also tackles the research gap identified in the matrix by Beattie, Murphy (16), indicating that the PEQ by Pettersen, Veenstra (15) lacked some validity analyses. This buttresses the claim that, indeed, changing theoretical underpinnings influences existing measures and changing statistical methods and tools can reveal weaknesses of measures; moreover, this should be countered by regular psychometric appraisals of these measures. The results also contribute to the views of some researchers (e.g. 17, 18), regarding the need to repeat patient experience surveys to generate more reliable data for policy-making. The assessment of patients' perspectives of hospital care would have to be reliable and valid enough in order to elicit accurate information about their experiences, constructs, and outcomes. Thus, it is imperative to ensure that these instruments always perform optimally and generate reliable information on how to improve quality of care and hospital experiences. These results, therefore, provide a background for further studies to be conducted on PREMs.

Another major contribution of this study is the finding of a second-order factor labelled *"treatment services"*, which consists of 4 factors: *"nurse services"*, *"doctor services"*, *"information"*, and *"organisation"*. This means that there exist strong and significant relationships among these dimensions (46). This finding also means that a single dimension or factor could adequately account for all 4 dimensions and could be identified as a major sub-dimension that captures these 4 dimensions. The *"treatment services"* factor has implications for the conceptualisation of patient-oriented hospital service climates. Patients in these hospitals may have highly overlapping experiences across *"nurse services"*, *"doctor services"*, *"organisation"*, and *"information"*. In more specific terms, it can be said that these patients experience a main dimension that accounts for significant portions of the 4 dimensions, perhaps because of the way these factors play out in the hospitals. For instance, doctors provide information regarding patients' health, ailments, and treatments while nurses organise and assist patients with the treatment process. This is significant in advancing knowledge of patient experiences. The experience of these 4 dimensions may not be that distinct, and patients, in experiencing service climate in the hospitals, may not adequately distinguish their shared perceptions of *"doctor services"* from *"information"* or of *"nurse services"* from *"organisation"*, for instance. Patient-oriented hospital service climate, as conceptualised in this study, may include 2 levels of factors that influence patients' perceptions and experiences of healthcare. Based on this, the definition of patient-oriented hospital service climate is modified as: patients' shared perceptions and experiences across different levels of hospital work environments, the services, and formal and informal practices and procedures that make up the entire caregiving process and that inform patient outcomes. This contribution is also a major highlight when

compared to the study by Pettersen, Veenstra (15), in which discriminant validity was not examined and a resulting second-order factor analysis was not explored. This challenges the theoretical structure of the PEQ and theoretical distinctness among these factors. Therefore, this study suggests a change in the factor structure of the PEQ and the development of a second-order factor for these 4 dimensions in the general patient experience literature. These possibilities are worth exploring in further surveys and studies on hospital factors as patient experiences during the caregiving process.

Practical Implications

Quality healthcare delivery is not exclusive to a region or country but a general goal of all healthcare systems worldwide. This can be contributed to by generating accurate information on how healthcare users experience healthcare systems. The results from this study suggest that it is not enough to develop a good measure of patient experiences, but it is imperative to review and reassess the ability of the measure to keep generating accurate information on patients' experiences and health. The questions in the PEQ may have to be revised in order to elicit more concise and accurate information from patients. Furthermore, some dimensions, such as *"next of kin"*, seemed not to be relatable to most of the patients, judging from the many missing values and invariance tests. Also, the PEQ should be administered with the second-order factor taken into consideration. It is imperative to analyse *"nurse services"*, *"doctor services"*, *"information"*, and *"organisation"* as a second-order factor, as shown in the proposed model, due to the validity issues that were realised in the analysis. This can provide researchers and management with adequate knowledge on what patients experience during the caregiving process. Moreover, management must take the interrelationships in the second-order factor into account to make meaningful, informed, and sustainable changes in the hospitals for patients. The second-order factor must be considered as a single climate encompassing these 4 dimensions, where patients' perceptions and interactions with a factor has a ripple effect on the others. Such considerations in policies and practice can help management and workers to reduce errors that may have dire consequences.

Limitations and directions for future research

This study employs data that is not at the national level but from a health region in Norway. That notwithstanding, the study has good generalisability power owing to the similarity in hospital and healthcare systems across the regions in Norway. Generalising to other countries, however, is difficult due to the differences in culture and healthcare systems. The findings require additional research in different countries for further justification. Therefore, future studies on reassessing psychometric properties of PREMs may want to employ larger data sets, for instance at the national level or across regions, to further investigate and develop the measurement quality of such surveys. Furthermore, future research should adopt the proposed model (with the second-order factor) from this study and examine it empirically to confirm it or otherwise, within health sectors across different countries.

Conclusion

Hospital management should know and consider the views and experiences of the people they care for if their services are to be influential in improving patients' health. The results of this study show that changes in psychometric analytical tools and methods can indeed highlight possible weaknesses and inadequacies in measures, as seen with the PEQ. This is evident in analyses such as invariance, discriminant validity, and second-order factors conducted in the current study but absent in the earlier study. Therefore, repeated surveys are needed to improve the measures' performance. The results also indicate possible changes with regard to dimensionality of PREMs, owing to the second-order factor finding. This calls for adequate attention, from researchers and hospital management alike, to the interrelationships among some of the dimensions, as this has important implications for theory and practice in healthcare. Management should consider these relationships in making decisions concerning the quality of care for patients, while researchers should delve more into studies that ascertain the psychometrics and dimensionality of PREMs.

List Of Abbreviations

PEQ- Patient Experience Questionnaire

PREM- Patient-reported Experience Measure

NIPH- The Norwegian Institute of Public Health

CFI- Comparative Fit Index

TLI- Tucker Lewis Index

RMSEA- Root Mean Square Error of Approximation

AVE- Average Variance Extracted

Declarations

- **Ethics approval and consent to participate**

This study, with regard to data collection, analysis, and compilation, was conducted within the ethical and legal provisions and guidelines of the Norwegian Institute for Public Health (NIPH) and the University of Stavanger. The Norwegian Data Protection Authority and the Norwegian Directorate of Health approved the procedures in the survey. The hospital data protection official assessed the data processing in the hospitals where survey extension took place. Informed consent was obtained from participants in the survey. Respondents were informed that participation was voluntary and they were assured of confidentiality of the information they will provide. Respondents were also informed that they could opt out of the survey at any point as well as the procedure for opting out if they wished. Data was stored in a safe repository with a password, only accessed by the researchers. This study did not present results that

revealed patients' identities, thus maintaining anonymity of respondents and confidentiality of responses. All relevant ethical requirements were duly upheld.

- **Consent for publication**

Not applicable

- **Availability of data and materials**

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

- **Competing interests**

The authors declare that they have no competing interests.

- **Funding**

Not applicable

- **Authors' contributions**

SA conducted the statistical analysis and the write-up. EO led the procurement of the data from NIPH, contributed to the statistical analysis and reviewed the write-up. RJM reviewed the write-up and contributed to the theoretical basis of the study. All authors read and approved the final manuscript.

- **Acknowledgements**

This study used anonymous data from the Norwegian Institute of Public Health. The interpretation and reporting of the data are the responsibility of the authors. We thank Ingeborg Strømseng Sjetne for preparing the data for this study and for her useful comments and suggestions.

References

1. Eng C-J, Pai H-C. Determinants of nursing competence of nursing students in Taiwan: The role of self-reflection and insight. *Nurse Education Today*. 2015;35(3):450-5.
2. Pitter JG, Csanádi M, Szigeti A, Lukács G, Kovács Á, Moizs M, et al. Planning, implementation and operation of a personalized patient management system for subjects with first suspect of cancer (OnkoNetwork): System description based on a qualitative study. *BMC Health Services Research*. 2019;19(1):131.
3. Wolf C, Jason A. Defining patient experience. *Patient Experience Journal*. 2014;1(1):7-19.
4. Glickman SW, Baggett KA, Krubert CG, Peterson ED, Schulman KA. Promoting quality: The health-care organization from a management perspective. *International Journal for Quality in Health Care*. 2007;19(6):341-8.

5. Lawson EF, Yazdany J. Healthcare quality in systemic lupus erythematosus: Using Donabedian's conceptual framework to understand what we know. *International Journal of Clinical Rheumatology*. 2012;7(1):95.
6. Netemeyer RG, Bearden WO, Sharma S. *Scaling procedures: Issues and applications*. Sage Publications; 2003.
7. DeVellis RF. *Scale development: Theory and applications*. Sage publications; 2016.
8. Bruyneel L, Tambuyzer E, Coeckelberghs E, De Wachter D, Sermeus W, De Ridder D, et al. New instrument to measure hospital patient experiences in Flanders. *International Journal of Environmental Research and Public Health*. 2017;14(11):1319.
9. Garratt AM, Bjertnæs ØA, Barlinn J. Parent experiences of paediatric care (PEPC) questionnaire: Reliability and validity following a national survey. *Acta Paediatrica*. 2007;96(2):246-52.
10. Garratt AM, Danielsen K, Forland O, Hunskaar S. The Patient Experiences Questionnaire for Out-of-Hours Care (PEQ-OHC): Data quality, reliability, and validity. *Scandinavian Journal of Primary Health Care*. 2010;28(2):95-101.
11. Iversen HH, Holmboe O, Bjertnæs ØA. The Cancer Patient Experiences Questionnaire (CPEQ): Reliability and construct validity following a national survey to assess hospital cancer care from the patient perspective. *BMJ open*. 2012;2(5):e001437.
12. Jenkinson C, Coulter A, Bruster S. The Picker Patient Experience Questionnaire: Development and validation using data from in-patient surveys in five countries. *International Journal for Quality in Health Care*. 2002;14(5):353-8.
13. Pettersen KI, Veenstra M, Guldvog B, Kolstad A. The Patient Experiences Questionnaire: Development, validity and reliability. *International Journal for Quality in Health Care*. 2004;16(6):453-63.
14. Beattie M, Murphy DJ, Atherton I, Lauder W. Instruments to measure patient experience of healthcare quality in hospitals: A systematic review. *Systematic Reviews*. 2015;4(1):97.
15. Haugum M, Danielsen K, Iversen HH, Bjertnaes O. The use of data from national and other large-scale user experience surveys in local quality work: A systematic review. *International Journal for Quality in Health Care*. 2014;26(6):592-605.
16. Manary MP, Boulding W, Staelin R, Glickman SW. The patient experience and health outcomes. *New England Journal of Medicine*. 2013;368(3):201-3.
17. Garratt A, Danielsen K, Bjertnaes Ø, Ruud T. PIPEQ—A method for measurement of user satisfaction in mental health services. *Tidsskrift for den Norske laegeforening: Tidsskrift for praktisk medicin, ny raekke*. 2006;126(11):1478-80.
18. Sjetne IS, Bjertnaes OA, Olsen RV, Iversen HH, Bukholm G. The Generic Short Patient Experiences Questionnaire (GS-PEQ): Identification of core items from a survey in Norway. *BMC Health Services Research*. 2011;11(1):88.
19. Ahern M, Dean CM, Dear BF, Willcock SM, Hush JM. The experiences and needs of people seeking primary care for low-back pain in Australia. *Pain Reports*. 2019;4(4).

20. Bjertnaes O, Deilkås ET, Skudal KE, Iversen HH, Bjerkan AM. The association between patient-reported incidents in hospitals and estimated rates of patient harm. *International Journal for Quality in Health Care*. 2014;27(1):26-30.
21. Bjertnaes OA, Sjetne IS, Iversen HH. Overall patient satisfaction with hospitals: Effects of patient-reported experiences and fulfilment of expectations. *BMJ Qual Saf*. 2012;21(1):39-46.
22. Bjorngaard JH, Garratt A, Grawe RW, Bjertnaes ØA, Ruud T. Patient experiences with treatment in private practice compared with public mental health services. *Scandinavian Journal of Psychology*. 2008;49(4):385-92.
23. Hinsley K, Kelly PJ, Davis E. Experiences of patient-centred care in alcohol and other drug treatment settings: A qualitative study to inform design of a patient-reported experience measure. *Drug and Alcohol Review*. 2019.
24. Litwin GH, Stringer RA. *Motivation and Organizational Climate*. 1968.
25. Reichers AE, Schneider B. Climate and culture: An evolution of constructs. *Organizational Climate and Culture*. 1990;1:5-39.
26. Denison DR. What is the difference between organizational culture and organizational climate? A native's point of view on a decade of paradigm wars. *Academy of Management Review*. 1996;21(3):619-54.
27. Zohar D. Safety climate in industrial organizations: Theoretical and applied implications. *Journal of Applied Psychology*. 1980;65(1):96.
28. Schneider B, Bowen DE. Employee and customer perceptions of service in banks: Replication and extension. *Journal of Applied Psychology*. 1985;70(3):423.
29. Schneider B, White SS, Paul MC. Linking service climate and customer perceptions of service quality: Tests of a causal model. *Journal of Applied Psychology*. 1998;83(2):150.
30. Speroff T, Nwosu S, Greevy R, Weinger M, Talbot T, Wall R, et al. Organisational culture: Variation across hospitals and connection to patient safety climate. *BMJ Quality & Safety*. 2010;19(6):592-6.
31. Bleich SN, Özaltın E, Murray CJ. How does satisfaction with the health-care system relate to patient experience? *Bulletin of the World Health Organization*. 2009;87:271-8.
32. Jenkinson C, Coulter A, Bruster S, Richards N, Chandola T. Patients' experiences and satisfaction with health care: Results of a questionnaire study of specific aspects of care. *Qual Saf Health Care*. 2002;11(4):335-9.
33. Taylor F, Halter M, Drennan VM. Understanding patients' satisfaction with physician assistant/associate encounters through communication experiences: A qualitative study in acute hospitals in England. *BMC Health Services Research*. 2019;19(1):603.
34. Hair J, Black W, Babin B, Anderson R, Tatham R. *Multivariate Data Analysis (6th Eds)*. New Jersey: Pearson Prentice Hall. 2006.
35. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*. 2006;99(6):323-

38.

36. Hair Jr JF, Sarstedt M, Hopkins L, Kuppelwieser VG. Partial least squares structural equation modeling (PLS-SEM). *European Business Review*. 2014.
37. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*. 1981;18(1):39-50.
38. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. Northridge, Cal: Harper Collins. 1996.
39. Vandenberg RJ, Lance CE. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*. 2000;3(1):4-70.
40. MacKenzie SB, Podsakoff PM, Podsakoff NP. Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*. 2011;35(2):293-334.
41. Gaskin J, Lim J. Master validity tool. AMOS Plugin In: Gaskination's StatWiki. 2016.
42. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*. 2016;41:71-90.
43. The OutPatient Experiences Questionnaire (OPEQ): Data quality, reliability, and validity in patients attending 52 Norwegian hospitals [Internet]. 2005.
44. Rindskopf D, Rose T. Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research*. 1988;23(1):51-67.
45. Skudal KE, Garratt AM, Eriksson B, Leinonen T, Simonsen J, Bjertnaes OA. The Nordic Patient Experiences Questionnaire (NORPEQ): Cross-national comparison of data quality, internal consistency and validity in four Nordic countries. *BMJ open*. 2012;2(3):e000864.

Tables

Table 1 Fitness indices and acceptable thresholds

| Fit indices | Acceptable thresholds |
|-------------|-------------------------------------|
| CFI | >.95, excellent; >.90, acceptable |
| TLI | >.95, excellent; >.90, acceptable |
| RMSEA | < .06, excellent; .06-.10, moderate |
| PCLOSE | > .05, excellent |

Adapted from Hu and Bentler (1999).

Table 2 Sample characteristics

| Variables | Frequency | Valid percent |
|------------------------|-----------|---------------|
| Age | | |
| Less than 61 years | 1502 | 32.6 |
| 61 - 73 years | 1528 | 33.2 |
| 73 years and above | 1573 | 34.2 |
| Total | 4603 | 100.0 |
| Days spent in hospital | | |
| Less than 4 days | 2630 | 57.1 |
| 4 or more days | 1973 | 42.9 |
| Total | 4603 | 100.0 |
| Department aggregates | | |
| Medical (Med) | 2468 | 53.6 |
| Surgical (Kir) | 2135 | 46.4 |
| Total | 4603 | 100.0 |
| Hospitals | | |
| Hospital 1 | 2067 | 44.9 |
| Hospital 2 | 1084 | 23.5 |
| Hospital 3 | 193 | 4.2 |
| Hospital 4 | 794 | 17.2 |
| Hospital 5 | 465 | 10.1 |
| Total | 4603 | 100.0 |

Table 3 Fitness results for all models

| Fit indices | Model 1- Initial model without modifications | Model 2- *Model after modifications | Models 3-8- *Alternative models | | | Model 9- *Configural invariance | Model 10- *Model after item deletion | Model 11- *Proposed model |
|-------------|--|-------------------------------------|------------------------------------|------------------------------------|------------------------------------|---------------------------------|--------------------------------------|---------------------------|
| | | | 1 st 4 th | 2 nd 5 th | 3 rd 6 th | | | |
| CFI | .91 | .95 | .91 .88 .85 | .91 .86 | .92 | .95 | .95 | .96 |
| TLI | .89 | .94 | .89 .85 .82 | .89 .84 | .90 | .94 | .94 | .95 |
| RMSEA | .06 | .04 | .06 .07 | .06 .07 | .05 .07 | .03 | .04 | .04 |
| PCLOSE | .00 | 1.00 | .00 .00 | .00 .00 | .00 .00 | 1.00 | .99 | .98 |

Note: *These models were assessed with the modification estimates. 1st- nurse and doctor into one factor; 2nd- nurse, doctor and organisation into one factor; 3rd- nurse and organisation into one factor; doctor and information into one factor; 4th- nurse, doctor, organisation and information into one factor; next of kin and standard into one factor; discharge and interaction into one factor; 5th- nurse, doctor, organisation, information, next of kin and standard into one factor; discharge and interaction into one factor; 6th- all dimensions into one factor.

Table 4 Correlations, reliability, convergent and discriminant validity before item deletion (Model 2)

| | CR | AVE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|--------------------|------|------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 1. Nurse services | 0.90 | 0.57 | 0.76 | | | | | | | |
| 2. Doctor services | 0.92 | 0.64 | 0.80 | 0.80 | | | | | | |
| 3. Information | 0.87 | 0.70 | 0.77 | 0.82 | 0.83 | | | | | |
| 4. Organisation | 0.81 | 0.53 | 0.85 | 0.80 | 0.79 | 0.73 | | | | |
| 5. Next of kin | 0.83 | 0.70 | 0.69 | 0.59 | 0.63 | 0.71 | 0.84 | | | |
| 6. Standard | 0.82 | 0.44 | 0.65 | 0.57 | 0.56 | 0.73 | 0.60 | 0.67 | | |
| 7. Discharge | 0.87 | 0.77 | 0.57 | 0.56 | 0.63 | 0.58 | 0.49 | 0.44 | 0.88 | |
| 8. Interaction | 0.72 | 0.57 | 0.57 | 0.58 | 0.56 | 0.63 | 0.53 | 0.50 | 0.56 | 0.75 |

Note: CR- Composite Reliability; AVE- Average Variance Explained; figures in bold are the square roots of the AVEs for discriminant validity (using the Fornell-Larcker procedure; discriminant validity is supported when the square root of the AVEs are greater than the correlation coefficients between the constructs).

Table 5 Standardised factor loadings (before and after item deletion) and missing values

| Dimensions and items | Factor loadings | | | |
|--|-----------------|----------|----------|----------------------|
| | Model 2 | Model 10 | Model 11 | Missing values N (%) |
| Nurse services | | | | |
| N1. Did the nursing staff talk to you so you understood them? | .67 | .66 | .67 | 287 (6.2) |
| N2. Did you find that the nursing staff cared for you? | .80 | .80 | .79 | 293 (6.4) |
| N3. Do you have confidence in the professional skills of the nursing staff? | .78 | .78 | .78 | 282 (6.2) |
| N4. Did you tell the nursing staff everything you thought was important about your condition? | .72 | .72 | .72 | 340 (7.4) |
| N5. Did you find that the nursing staff were interested in your description of your own situation? | .83 | .83 | .83 | 328 (7.1) |
| N6. Were you included in the advice on questions regarding your care? | .70 | .70 | .68 | 427 (9.3) |
| N7. Did the nursing staff have time for you when you needed it? | .77 | .77 | .78 | 297 (6.5) |
| Doctor services | | | | |
| D1. Did the doctors talk to you so you understood them? | .73 | .73 | .73 | 300 (6.5) |
| D2. Did you find that the doctors took care of you? | .84 | .84 | .83 | 302 (6.6) |
| D3. Do you trust the doctors' professional skills? | .77 | .77 | .77 | 299 (6.5) |
| D4. Did the doctors have time for you when you needed it? | .82 | .82 | .82 | 415 (9.0) |
| D5. Did you tell the doctors everything you thought was important about your condition? | .77 | .77 | .77 | 384 (8.3) |
| D6. Did you find that the doctors were interested in your description of your own situation? | .84 | .84 | .85 | 378 (8.2) |
| D7. Did you find that the treatment was adapted to your situation? | .79 | .79 | .76 | 321 (7.0) |
| Information | | | | |
| IF1. Did you know what you thought was necessary about how tests and examinations should take place? | .79 | .79 | .85 | 320 (7.0) |
| IF2. Did you know what you thought was necessary about the results of tests and examinations? | .85 | .85 | .86 | 334 (7.3) |
| IF3. Did you receive sufficient information about your diagnosis or your complaints? | .86 | .86 | .87 | 326 (7.1) |
| Dimensions and items | Factor loadings | | | |
| | Model 2 | Model 10 | Model 11 | Missing values N (%) |
| Organisation | | | | |
| OR1. Did you find that there was a permanent group of nursing staff that took care of you? | .67 | .67 | .68 | 121 (2.6) |
| *OR2. Did you find that one doctor had the main responsibility for you? | .58 | - | - | 130 (2.8) |
| OR3. Did you find that the hospital's work was well organized? | .81 | .82 | .82 | 107 (2.3) |
| OR4. Did you find that important information about you had come to the right person? | .82 | .81 | .81 | 204 (4.4) |
| Next of kin | | | | |
| NK1. Were your relatives well received by the hospital staff? | .84 | .84 | .84 | 1362 (29.6) |
| NK2. Was it easy for your relatives to get information about you while you were in the hospital? | .83 | .83 | .83 | 1732 (37.6) |

| | | | | |
|--|---------|-----------------|----------|----------------|
| Standard | | | | |
| S1. Did you get the impression that the hospital equipment was in good condition? | .71 | .72 | .78 | 108 (2.3) |
| S2. Did you get the impression that the hospital was in good condition? | .77 | .78 | .86 | 122 (2.7) |
| S3. Was the room you were in satisfactory? | .67 | .65 | .74 | 80 (1.7) |
| *S4. Was the opportunity for rest and rest satisfactory? | .62 | - | .66 | 90 (2.0) |
| *S5. Was the food satisfactory? | .55 | - | - | 122 (2.7) |
| S6. Was the cleaning satisfactory? | .65 | .66 | .60 | 89 (1.9) |
| Discharge | | | | |
| DC1. Were you informed of what you could do at home in case of relapse? | .87 | .87 | - | 1327 (28.8) |
| DC2. Were you informed of what complaints you could expect to receive in time after your hospital stay? | .88 | .88 | - | 1195 (26.0) |
| Interaction | | | | |
| IT1. Do you find that the hospital has worked well with your GP about what you were admitted to? | .82 | .81 | - | 2523 (33.9) |
| Dimensions and items | | Factor loadings | | |
| | Model 2 | Model 10 | Model 11 | |
| IT2. Do you feel that the hospital has cooperated well with the home or other municipal services about what you were admitted for? | .69 | .69 | - | 3401 (54.8) |
| Treatment services | | | | |
| Nurse services | | | | .92 |
| Doctor services | | | | .86 |
| Information | | | | .84 |
| Organisation | | | | .93 |

Note: Items marked with * had the lowest loadings

Table 6 Correlations, reliability, convergent and discriminant validity after item deletion (Model 10)

| | CR | AVE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|--------------------|------|------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 1. Nurse services | 0.90 | 0.57 | 0.76 | | | | | | | |
| 2. Doctor services | 0.92 | 0.64 | 0.80 | 0.80 | | | | | | |
| 3. Information | 0.87 | 0.70 | 0.77 | 0.82 | 0.83 | | | | | |
| 4. Organisation | 0.81 | 0.59 | 0.87 | 0.78 | 0.79 | 0.77 | | | | |
| 5. Next of kin | 0.82 | 0.70 | 0.69 | 0.59 | 0.63 | 0.72 | 0.84 | | | |
| 6. Standard | 0.80 | 0.50 | 0.64 | 0.56 | 0.55 | 0.74 | 0.59 | 0.70 | | |
| 7. Discharge | 0.87 | 0.77 | 0.57 | 0.56 | 0.63 | 0.56 | 0.49 | 0.44 | 0.88 | |
| 8. Interaction | 0.72 | 0.57 | 0.57 | 0.58 | 0.56 | 0.62 | 0.53 | 0.47 | 0.56 | 0.75 |

Note: CR- Composite Reliability; AVE- Average Variance Explained; figures in bold are the square roots of the AVEs for discriminant validity (using the Fornell-Larcker procedure; discriminant validity is supported when the square root of the AVEs are greater than the correlation coefficients between the constructs).

Table 7 Regression results for criterion-related validity

| | | | | Outcome variables | | | |
|--|--|--|--|---------------------------|--|--|--|
| | | | | Proposed model (Model 11) | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |

***p<0.001; **p<.01; *p<.05; empty fields are not significant at 0.05 level; Treatment services – second order factor comprising nurse services, doctor services, information and organisation.

Table 8 Tools and findings in the earlier validation study and the current study

| Study | Psychometric tools used | Findings |
|-------------------------|---|--|
| Pettersen et al. (2004) | Exploratory factor analysis Cronbach's alpha test Test-retest reliability Construct validity | 10 factors (including general satisfaction) Confirmed Confirmed Achieved |
| Current study | Confirmatory factor analysis Model comparisons Measurement invariance Composite reliability test Convergent validity Discriminant validity Construct validity Criterion-related validity Second-order factor analysis | 8 factors (excluding general satisfaction) Initial model was found to be best Configural and Metric achieved, Scalar not achieved Confirmed Confirmed for all except one factor Confirmed for all except three factors Achieved Achieved Achieved composite reliability, convergent validity, discriminant validity, construct validity and criterion related validity |

Figures

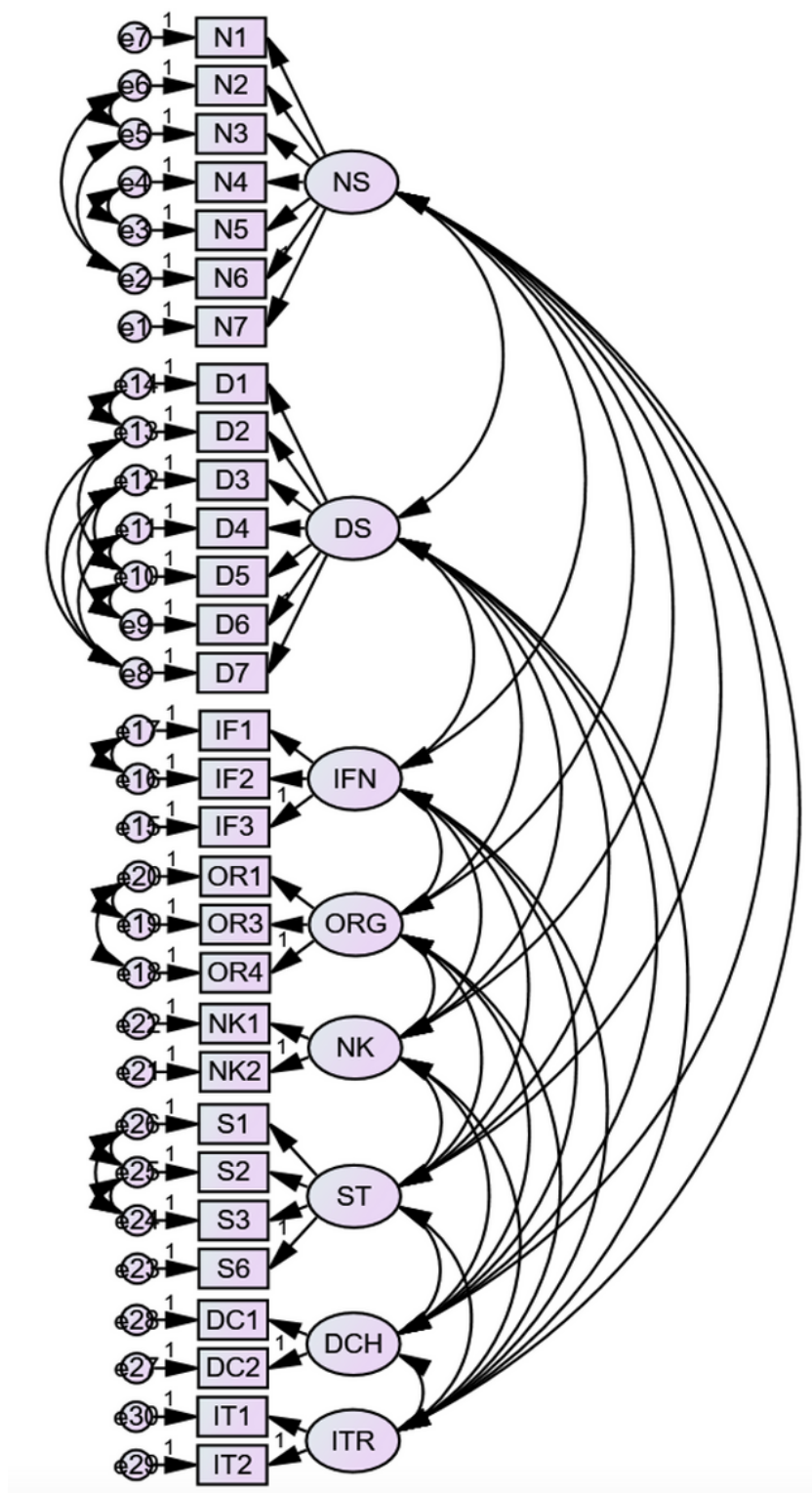


Figure 1

Model after validity checks and item deletion. Note- NS: nurse services; DS: doctor services; INF: information; ORG: organisation; ST: standard; NK: next of kin; DCH: discharge; ITR: interaction

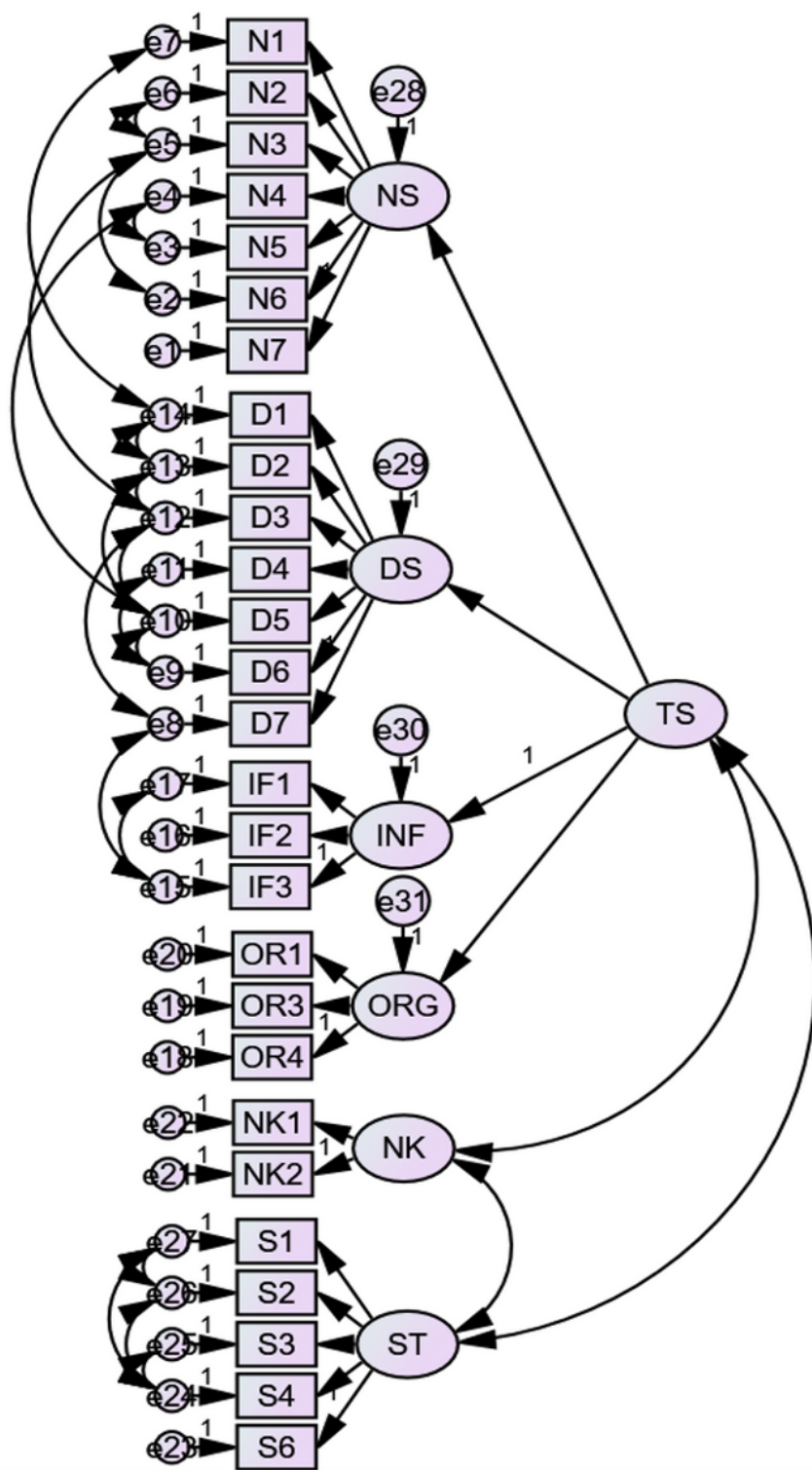


Figure 2

Proposed measurement model

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarytables.docx](#)