Combining principal component analysis and the evidential reasoning approach for healthcare quality assessment

DOI:
10.1007/s10479-018-2789-z

Document Version
Accepted author manuscript

Citation for published version (APA):
Kong, G., Jiang, L., Yin, X., Wang, T., Xu, D.-L., Yang, J.-B., & Hu, Y. (2018). Combining principal component analysis and the evidential reasoning approach for healthcare quality assessment. *Annals of Operations Research*. https://doi.org/10.1007/s10479-018-2789-z

Published in:
Annals of Operations Research

Combining principal component analysis and the evidential reasoning approach for healthcare quality assessment

Guilan Kong[a,*], Lili Jiang[b], Xiaofeng Yin[c], Tianbing Wang[c],

Dong-Ling Xu[d], Jian-Bo Yang[d], Yonghua Hu[a,e]

[a*]: Medical Informatics Center, Peking University, Beijing, 100191, China

Tel.: +86 10 82805902. E-mail address: Guilan.kong@hsc.pku.edu.cn

[b]: Taikang Life Insurance Co. Ltd,, Beijing, 100031, China

[c]: Department of Trauma and Orthopaedics, Peking University People's Hospital, Beijing, 100044, China

[d]: Decision and Cognitive Sciences Research Centre, The University of Manchester, Manchester M15 6PB, UK

[e]:School of Public Health, Peking University, Beijing, 100191, China.

Abstract

Patient experience and satisfaction surveys have been adopted worldwide to evaluate healthcare quality. Nevertheless, national governments and the general public continue to search for optimal methods to assess healthcare quality from the patient's perspective. This study proposes a new hybrid method, which combines principal component analysis (PCA) and the evidential reasoning (ER) approach, for assessing patient satisfaction. PCA is utilized to transform correlated items into a few uncorrelated principal components (PCs). Then, the ER approach is employed to aggregate extracted PCs, which are considered as multiple attributes or criteria within the ER framework. To compare the performance of the proposed method with that of another assessment method, analytic hierarchy process (AHP) is employed to acquire the weight of each assessment item in the hierarchical assessment framework, and the ER approach is used to aggregate patient evaluation for each item. Compared with the combined AHP and ER approach, which relies on the respondents' subjective judgments to calculate criterion and subcriterion weights in the assessment framework, the proposed method is highly objective and completely based on survey data. This study contributes a novel and innovative hybrid method that can help hospital administrators obtain an objective and aggregated healthcare quality assessment based on patient experience.

Keywords: Healthcare quality assessment, patient experience and satisfaction, principal component analysis, analytic hierarchy process, the evidential reasoning approach

1. Background

Healthcare quality assessment has become a crucial topic of healthcare studies given that it will help ensure the proper allocation of limited healthcare resources in the face of continuously increasing healthcare demands and costs and standardize medical practice (Büyüközkan & Çifçi 2012, Büyüközkan et al 2011, Fragkiadakis et al 2016, Kong et al 2015, Lyratzopoulos et al 2011, Panagiotis et al 2016, Prior 2006). Patient experience is an important healthcare outcome, and surveys that measure patient experience and satisfaction are currently widely used to assess healthcare quality (Department of Health 2013, Jenkinson et al 2002, Jha et al 2008, Keller et al 2005, Kleefstra et al 2010, Vuković et al 2012). Governments and regulatory authorities in some countries now require hospitals to organize patient surveys at regular intervals (Jenkinson et al 2002). In the United States (US), some policy initiatives have attached financial incentives, such as directly linking patients' evaluations with doctors' financial rewards, to patient surveys (Rodriguez et al 2009). In the United Kingdom (UK), the Department of Health (2000) has launched a program of national surveys and has required every National Health Service (NHS) Trust to survey their patients annually. In Switzerland, the National Coordination and Information Office for Quality Improvement has recommended that a survey instrument be administered to hundreds of hospitals on an annual basis (Jenkinson et al 2002). In Australia, a national patient experience survey is conducted annually (Department of Health 2013). In China, the national government launched a new wave of healthcare reform in 2009 to reduce healthcare costs and improve healthcare quality and patient safety. To achieve these goals, the current healthcare strategy in China links the healthcare quality of hospitals with the allocation of healthcare resources, such as government funding. The National Health and Family Planning Commission of China requires that a patient experience survey be an integral component of healthcare quality assessment.

A review of the literature shows that in different countries and regions, different questionnaires have been used to measure healthcare quality from different dimensions. In the US, the Centers for Medicare and Medicaid Services has collaborated with the Agency for Healthcare Research and Quality to develop a standardized patient satisfaction questionnaire, the Consumer Assessment of Health Providers and Systems, for measuring the quality of inpatient hospital care (Goldstein et al 2005, Jha et al 2008). In the UK, the Picker Patient Experience questionnaire is used to measure patients' experiences of inpatient care (Jenkinson et al 2003, Keller et al 2014). This questionnaire is given annually to survey the quality of inpatient care provided by all hospitals belonging to the NHS system. Moreover, since 2000, the Department of Health has required that the results of the survey must be reported in an annual patient prospectus. Until 2013, the Victoria Patient Satisfaction Monitor was the most widely used inpatient satisfaction questionnaire in Australia. This questionnaire has now been replaced by the Victorian Health Experience Measurement Instrument (Department of Health 2013). In the Netherlands, eight academic hospitals have developed a Core Questionnaire for the Assessment of Patient Satisfaction (COPS) (Kleefstra et al 2010). The Federation of Dutch Hospitals has accepted COPS as a standard instrument for measuring patient satisfaction. The main healthcare dimensions measured by the above questionnaires include: doctor–patient or nurse–patient communication; staff responsiveness; environmental cleanliness and noise level; pain control or physical comfort; drug, admission, or discharge information communication; and overall satisfaction.

Different statistical methods have been employed to analyze survey data for patient experience. Spearman correlation analysis has been used to analyze the relationships between survey items and overall evaluation (Jenkinson et al 2002, Keller et al 2014). Cronbach's α coefficient has been used to measure the internal consistency and reliability of questionnaires (Harris et al 1999, Keller et al 2005, Purcărea et al 2013, Vuković et al 2012). Exploratory and confirmatory factor analyses have been used to explore and validate the structure of the measured dimensions and items of questionnaires (Harris et al 1999, Keller et al 2014, Keller et al 2005). Regression models have been used to determine the impact of individual items on overall quality evaluation (Vuković et al 2012, Wong et al 2011). Multidimensional scaling has been used to identify similarities and dissimilarities among items in questionnaires (Vuković et al 2012), and principal component analysis (PCA) has been used to identify the main healthcare dimensions and their relationships with individual measured items from survey data (Purcărea et al 2013, Vuković et al 2012).

However, all the above statistical methods are for questionnaire validation or total-item relationship exploration, and advanced decision models that combine patient assessments or evaluations of different items or variables are needed to measure and evaluate overall healthcare quality. Driven by the need for the combined or integrated assessment of overall healthcare quality, Behara et al. (2002) and Carlucci et al. (2013) used an artificial neural network (ANN) to model and obtain an overall evaluation from patient assessments of different healthcare dimensions. Büyüközkan et al. (2011) extended the traditional analytic hierarchy process (AHP) methodology to a fuzzy AHP to combine subjective and vague judgments of multiple healthcare quality indices or items. Büyüközkan and Çifçi (2012) combined a fuzzy AHP and a fuzzy technique for order of preference by similarity to ideal solution (TOPSIS) to aggregate patient assessments of multiple quality items. However, these combined assessment methods have their shortcomings. Specifically, an ANN contains nonlinear functions and is a black-box for users; these characteristics complicate its adoption by healthcare practitioners. Although the fuzzy AHP method extends the traditional AHP method to vague subjective judgments of multiple criteria and has the advantage of converting subjective judgments to numerical values, it contains the problem of rank reversal. Similar to the fuzzy AHP method, the fuzzy TOPSIS method has the advantage of handling fuzzy judgments of multiple criteria and the problem of rank reversal, which means that the ranking of alternatives may change when new alternatives are added. In our previous study (Kong et al 2015), we proposed using the evidential reasoning (ER) approach (Wang et al 2006, Xu 2012, Yang & Singh 1994) to combine objective quality indicators, subjective expert judgments, and patient feedback to provide an overall assessment of healthcare quality. The ER approach requires that the items measured or assessed in a questionnaire should be uncorrelated if their assessments are combined to obtain an overall quality assessment. In our previous study, we considered that patient evaluations on four items—medical facilities, medical staff, medical processes, and medical outcomes in a hospital—are independent of each other. Thus, the ER approach is suitable for combining patient assessments of these four items. However, most patient experience surveys for measuring healthcare quality include dozens of items, and some items are correlated to some degree. In this situation, applying the ER approach directly to combine assessments of individual items to obtain an overall quality assessment is irrational.

In the present study, we propose combining PCA (Jolliffe 2002) and the ER approach to aggregate patient assessments of multiple correlated items for overall healthcare quality assessment. PCA helps transform original interrelated variables into a new set of uncorrelated variables, the new principal components (PCs). The weights of these PCs are then determined in accordance with the amount of variance that each PC accounts for in the dataset. The weighted uncorrelated PCs are then used as new quality criterion variables and are combined through the ER approach to obtain an overall healthcare quality assessment. Meanwhile, to compare the performance of the proposed method with that of another method, AHP is employed to acquire the weights of different healthcare quality dimensions and their corresponding survey items, and. the ER approach is then used to aggregate the patient evaluation of each item.

The rest of this paper is structured as follows. The materials and methods used in this study are discussed in Section 2. The questionnaire is introduced in Section 2.1. The collected survey data are briefly discussed in Section 2.2. Brief introductions to

PCA, AHP, and the ER approach are provided in Sections 2.3, 2.4, and 2.5, respectively. The combined PCA and ER approach for the aggregation of patient assessments is introduced in Section 2.6. The combined AHP and ER approach for the aggregation of patient evaluations is described in Section 2.7. In addition to the characteristics of the survey data, the extracted PCs together with corresponding observable variables or items with significant component loadings, the weight of each extracted PC, the weights of different quality dimensions and corresponding survey items calculated via AHP, and the overall quality assessment results of both methods are presented in Section 3. Finally, a summary of this study and a discussion of the findings is provided in Section 4.

2. Materials and methods

2.1. Questionnaire

We developed a questionnaire in reference to survey instruments for patient experience that have been used in the UK, the US, the Netherlands, and Australia. In addition to demographic information about the respondents, the questionnaire provides one overall rating of healthcare quality. It contains 25 items that measure healthcare quality from various aspects or dimensions, such as hospital environment, waiting time, communication with doctors, communication with nurses, care coordination, physical comfort, emotional support, respect for patient preferences, family and friend involvement, and drug information. For each item, typical five-point Likert-type scale responses (“very dissatisfied,” “dissatisfied,” “fair,” “satisfied,” and “very satisfied,”) were adopted. Occasionally, “not applicable” was recorded by the researchers if the patients did not experience the problem associated with the question item. We coded "not applicable" responses as missing. For the overall rating of the healthcare quality, the satisfaction score of 0–10 was applied, where a score of 10 refers to the highest level of satisfaction.

2.2. Dataset

Between August and September 2014, all patients at the point of discharge from one department of a top-tier teaching hospital affiliated with Peking University (hereafter referred to as Hospital A), Beijing, China, received questionnaires assessing the quality of the healthcare they received on the basis of on their in-hospital experiences. All questionnaires were completed anonymously, and one of our researchers helped respondents eliminate worries about the consequences of their responses and provided instructions on answering the questionnaires. A total of 213 surveys were collected from the hospital. We did not send questionnaires to patients who were unwilling to give us their responses or assessments of received healthcare.

We preprocessed the data from the 213 collected surveys as follows. First, if the response rate for an item was lower than 90%, we excluded the item from data analysis. Second, we excluded a patient's survey data from the analysis if his or her responses to two or more items were "not applicable." Third, we calculated the median value for each item and used the median value to replace the missing data of items retained for analysis. Fourth, we employed Spearman correlation analysis to explore the item-total relationship and excluded items with correlation coefficients with values less than 0.3.

After data preprocessing, we obtained 192 valid surveys with six deleted items.

2.3. PCA

PCA is a multivariate statistical approach commonly used to reduce the dimensions of a dataset that consists of interrelated single indicators or variables. It is a linear combination of variables that explains the variance structure of a matrix and reduces various data into a few PCs. It focuses on the use of a few PCs to reveal the internal structure among multiple observable variables that are uncorrelated with each other and allows the PCs to preserve the information embodied in original variables as much as possible.

Let x be a vector of p random variables, and the variances of p variables and structures of the covariances or correlations between p variables are considered of interest. Consider X is a ($n \times p$) matrix with n observations on p variables, and K is the covariance matrix of the random vector x with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, and eigenvectors $\alpha_1, \alpha_2, \cdots, \alpha_p$. PCs are derived from the X matrix with the following linear functions $\alpha_j' x \ (j = 1,2 \cdots, p)$ of the elements of x, and the extracted PCs have maximum variance with constraints of $\alpha_j' x$ being uncorrelated, i.e., $Cov[\alpha_i' x, \alpha_j' x] = 0, (i \neq j)$ (Jolliffe 2002, Park et al 2015). The mathematical framework of PCA is as follows:

$$Z_1 = \alpha_1' x = \alpha_{11} x_1 + \alpha_{12} x_2 + \cdots + \alpha_{1p} x_p = \sum_{j=1}^{p} \alpha_{1j} x_j \quad (1)$$

$$Z_2 = \alpha_2' x = \alpha_{21} x_1 + \alpha_{22} x_2 + \cdots + \alpha_{2p} x_p = \sum_{j=1}^{p} \alpha_{2j} x_j \quad (2)$$

$$\vdots$$

$$Z_p = \alpha_p' x = \alpha_{p1} x_1 + \alpha_{p2} x_2 + \cdots + \alpha_{pp} x_p = \sum_{j=1}^{p} \alpha_{pj} x_j \quad (3)$$

$$Var[Z_i] = \alpha_i' \mathrm{K} \alpha_i, i = 1,2,\cdots,p \quad (4)$$

$$Cov[Z_i, Z_j] = \alpha_i' \mathrm{K} \alpha_j, i = 1,2,\cdots,p;\ j = 1,2,\cdots,p \quad (5)$$

where α_j is a vector of p coefficients $\alpha_{j1}, \alpha_{j2}, \cdots, \alpha_{jp}$, and α_j is nothing but the eigenvector of covariance matrix K that corresponds to the *j*th largest eigenvalue λ_j. $Z_i (i = 1,2,\cdots,p)$ represents PCs and '′' represents the transposition operation. The first linear function $\alpha_1' x$ finds the first PC, Z_1, that accounts for the maximal amount of total variance in the dataset. The second PC, Z_2, is uncorrelated with Z_1 and accounts for the maximal amount of variance in the dataset that is not accounted for by the first component, such that at the *k*th stage, a linear function $\alpha_k' x$ is found that has maximum variance subject to being uncorrelated with $\alpha_1' x, \alpha_2' x, \cdots, \alpha_{k-1}' x$. The *k*th derived variable, $\alpha_k' x$, is the *k*th PC. Up to *p* PCs can be found but in general most of the variation in *x* will be accounted for by *m* PCs where $m \leq p$. The elements in the diagonal of the covariance matrix of the derived PCs are known as the eigenvalues $\lambda_i (i = 1,2,\cdots,p)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, which are the variance explained by each PC and are constrained to decrease monotonically from the first PC to the last. The coefficient $\alpha_{ij} (i = 1,2,\cdots p; j = 1,2,\cdots,p)$ is the element of the eigenvector and is known as the loading or weight of the *j*th original variable for the ith PC (Jolliffe 2002). The importance or weight of each PC can be determined on the basis of the amount of variance that it accounts for in the dataset.

After extracting PCs from the original dataset through linear transformation, we need to understand the extracted PCs or determine which variables load significantly on which component to retain only loadings that are statistically significant for each

PC. Thus, we have to identify which variable loadings are significant and which can be safely ignored for each component. Usually, rotating the extracted components can help identify the variables that load strongly on each component (Norman & Streiner 1998). Therefore, the value or score of the extracted PCs can be computed from original variables by multiplying the standardized values of variables by their corresponding weights or coefficients. Sometimes, the values of extracted PCs can be computed only from variables with significant loadings (Norman & Streiner 1998).

2.4. AHP

AHP was first developed in 1971 by Thomas Saaty (Saaty 1980). It is a multicriterion decision analysis method in which a complex, multicriterion problem is decomposed into multiple levels of hierarchy with the top level as the goal, intermediate levels as the criteria and subcriteria, and the lowest level offers alternatives; a hierarchal structure is thus formed for assessment (Saaty 1980). The relative importance of all criteria and subcriteria within each level of hierarchy is usually determined by expert judgment and calculated through pairwise comparisons (Saaty 2008).

The typical application of AHP includes four main stages. First, a hierarchy of criteria used for assessment needs to be developed. Second, a pairwise comparison survey is conducted to elicit the preferences of respondents. At this stage, a pairwise comparison matrix is formed where w_i/w_j measures the importance of criterion i relative to j. Typically, a nine-point scale is used where 1 means equal importance between two criteria, and 9 means the extreme importance of one criterion compared with another. Third, the consistency of respondents' judgments in pairwise comparison is checked. Numerous methods, such as Eigenvalue method and geometric mean, are used to calculate the normalized weights of each criterion (Morgan 2017). In this study, we employed the Eigenvalue method for calculation. In the Eigenvalue method, a consistency ratio (CR) is employed to measure the consistency of individual responses, where 0 means perfect consistency in the responses given by an respondent and a CR value of 10% or less indicates that the pairwise comparison matrix is acceptable (Ishizaka et al 2010). Finally, the relative importance of each criterion in the hierarchy is calculated.

2.5. The ER approach

The ER approach (Xu 2012, Yang & Singh 1994, Yang & Xu 2002) was originally proposed to aid multiple attribute decision analysis (MADA) problems. It has the advantage of dealing with qualitative and quantitative attributes under uncertainty (Yang 2001, Yang & Xu 2002). It has been employed to aid medical decision-making, such as the assessment of clinical risk associated with cardiac chest pain (Kong et al 2009, Kong et al 2012) and combined healthcare quality assessment(Kong et al 2015).

We assume N alternatives $D_1, D_2, \cdots, D_N$ exist that need to be assessed on the basis of L individual attributes or indicators $A(A_1, A_2, \cdots, A_L)$, which are uncorrelated. The jth attribute $A_j (j = 1,2, \cdots, L)$ can either be qualitative or quantitative, and each attribute A_j can be assessed through a set of assessment grades $H(H_1, H_2, \cdots, H_M)$, which are assumed to be collectively exhaustive and mutually exclusive. Instead of using a certain score that represents an assessment grade to denote the evaluation of an alternative on an individual indicator in conventional MADA methods, a belief distribution, such as $\{(\beta_1, H_1), (\beta_2, H_2), \cdots, (\beta_M, H_M)\}$, can be used to express an evaluation of an indicator that is distributed on a fixed set of assessment grades *H*. Considering the relative importance or weight $\omega_j (j = 1,2, \cdots, L)$ of each measured attribute or indicator, a MADA problem can be modeled by the ER approach, as shown in Fig. 1, where $\beta_{tj} (t = 1,2, \cdots, M; j = 1,2, \cdots, L)$ is used to denote the degree of belief in the *t*th assessment grade H_t for assessing the *j*th attribute A_j. The belief degree can either be subjective if it quantifies a "personal belief" or objective if it is a computed probability on the basis of recorded data.

INSERT FIGURE 1 HERE

The core of the ER approach is the ER algorithm, which is used to aggregate the distributed assessments of all attributes or indicators and generate a combined assessment of an alternative. A brief introduction to the ER algorithm is provided below.

First of all, the degrees of belief $\beta_{tj} (t = 1,2, \cdots, M; j = 1,2, \cdots, L)$ are transformed into basic probability masses by combining the relative weights and the degrees of belief using the following equations:

$$m_{t,j} = w_j \beta_{tj}, t = 1,2, \cdots, M; j = 1,2, \cdots, L \tag{6}$$

$$m_{H,j} = 1 - \sum_{t=1}^{M} m_{t,j} = 1 - w_j \sum_{t=1}^{M} \beta_{tj}, j = 1,2, \cdots, L \tag{7}$$

$$\bar{m}_{H,j} = 1 - w_j, j = 1,2, \cdots, L \tag{8}$$

$$\tilde{m}_{H,j} = w_j \left(1 - \sum_{t=1}^{M} \beta_{tj}\right), j = 1,2, \cdots, L \tag{9}$$

where $m_{H,j} = \bar{m}_{H,j} + \tilde{m}_{H,j}$ for all j = 1,2,…,L and $\sum_{j=1}^{L} w_j = 1$. $m_{t,j}$ represents the basic probability mass of an alternative being assessed to the assessment grade H_t on attribute A_j. Note that the probability mass assigned to the grade set H, $m_{H,j}$, which is currently unassigned to any individual grades, is split into two parts: $\bar{m}_{H,j}$ and $\tilde{m}_{H,j}$. $\bar{m}_{H,j}$ is caused by the relative importance of the *j*th attribute A_j and $\tilde{m}_{H,j}$ is caused by the incompleteness of the *j*th attribute A_j. $\bar{m}_{H,j}$ represents the contribution of other attributes to assessing an alternative and is the proportion of beliefs that remain to be assigned in accordance with the assessment of other attributes. In essence, $\bar{m}_{H,j}$ provides a scope for conflict resolution in the presence of conflicting evidence. $\tilde{m}_{H,j}$ will be zero if ignorance is absent from the assessment.

Subsequently, all the distributed assessments on L attributes or indicators are aggregated to generate the combined degree of belief in each possible grade H_t. The analytic format of the ER aggregation algorithm (Wang et al 2006) is as follows:

$$m_t = k\left[\prod_{j=1}^{L}\left(m_{t,j} + \bar{m}_{H,j} + \tilde{m}_{H,j}\right) - \prod_{j=1}^{L}\left(\bar{m}_{H,j} + \tilde{m}_{H,j}\right)\right], t = 1,2,\cdots,M \quad (10)$$

$$\tilde{m}_H = k\left[\prod_{j=1}^{L}\left(\bar{m}_{H,j} + \tilde{m}_{H,j}\right) - \prod_{j=1}^{L}\bar{m}_{H,j}\right] \quad (11)$$

$$\bar{m}_H = k\left[\prod_{j=1}^{L}\bar{m}_{H,j}\right] \quad (12)$$

$$k = \left[\sum_{t=1}^{M}\prod_{j=1}^{L}\left(m_{t,j} + \bar{m}_{H,j} + \tilde{m}_{H,j}\right) - (M-1)\prod_{j=1}^{L}\left(\bar{m}_{H,j} + \tilde{m}_{H,j}\right)\right]^{-1} \quad (13)$$

$$\beta_t = \frac{m_t}{1-\bar{m}_H}, t = 1,2,\cdots,M \quad (14)$$

$$\beta_H = \frac{\tilde{m}_H}{1-\bar{m}_H} \quad (15)$$

where β_t and β_H represent the belief degrees of the aggregated assessment to which an alternative is assessed to grade H_t and H, respectively, after combining the distributed assessments on all indicators. The combined assessment of an alternative can be denoted by $S(y) = \{(H_t, \beta_t), t = 1,2,\cdots,M\}$. $\sum_{t=1}^{M}\beta_t + \beta_H = 1$ has been proven (Yang & Xu 2002).

2.6. Combining PCA and the ER approach to assess healthcare quality

As discussed in Sections 2.3 and 2.5, PCA has the advantage of transforming multiple interrelated indicators into a few uncorrelated PCs, and the ER approach has the advantage of combining the distributed assessments of multiple uncorrelated indicators under uncertainty. The combined PCA and ER approach can help rationally use collected survey data to provide an objective and aggregated healthcare quality assessment based on patient experience. The detailed procedures for combining PCA with the ER approach to assess the quality of healthcare provided by Hospital A are as follows:

First, numerical scores are used to replace the five-point Likert-type scales used in the survey. Specifically, a value of 1 is assigned to “very dissatisfied,” 2 to “dissatisfied,” 3 to “fair,” 4 to “satisfied,” and 5 to “very satisfied.” In this study, we obtained a numerical matrix $A(192 \times 19)$ after excluding unqualified patient surveys, and each item $a_{ij}(i = 1,2,\cdots, 192; j = 1,2,\cdots, 19)$ in the matrix ranges from 1 to 5.

Second, a preliminary statistical test, the Kaiser–Meyer–Olkin (KMO) index, accompanied by Bartlett’s test of Sphericity, should be employed to examine whether items in the survey dataset are interrelated. Moreover, the KMO test must have values higher than 0.5 and Bartlett’s test must be significant at a level lower than 0.05 (Purcărea et al 2013).

Third, if the survey dataset is suitable for PCA, PCA can be used to analyze the dataset and derive the PCs that can be used as uncorrelated criterion variables for an aggregated quality assessment. We employed SPSS software to perform PCA. SPSS provides two options for performing PCA: “correlation matrix” and “covariance matrix.” The default setting is “correlation matrix,” and we usually use the default “correlation matrix” to perform PCA. Nevertheless, if the original dataset has been standardized, performing PCA with the “covariance matrix” will yield the same results as the “correlation matrix”.

Fourth, PCs are extracted from PCA. Generally, three methods are used to extract PCs. One method is based on the eigenvalue of each PC, and PCs with eigenvalues larger than 1 can be extracted as final PCs for subsequent analysis. One method is based on the researchers’ subjective judgments of the number of PCs that need to be extracted. Thus, a fixed number of PCs can be extracted. Another method to determine the number of PCs that can be extracted is based on the cumulative variance for which all extracted PCs can account for. In this method, a threshold value is set for the cumulative variance proportion, and the number of PCs can then be determined if the cumulative variance of combined PCs has reached this threshold value. In this study, we set the threshold value of the cumulative variance proportion at 70%.

Fifth, weights that correspond to the extracted PCs are calculated for later aggregation using the ER approach. In our case, we employed the eigenvalues that correspond to the extracted PCs to calculate the weight of each PC. Given that only a proportion of PCs have been extracted to represent all the original surveyed items, we normalized the eigenvalues of the extracted PCs to obtain the weights of the corresponding PCs for later assessment aggregation. Assuming that m PCs have been extracted, and the corresponding eigenvalues are $\lambda_i (i = 1,2,\cdots,m)$, the weight associated with each extracted PC is calculated using the following:

$$w_i = \frac{\lambda_i}{\sum_{i=1}^{m} \lambda_i}, (i = 1,2,\cdots,m) \tag{16}$$

Sixth, variables that strongly load on each extracted PC are discovered, and the assessments distributed on different evaluation grades for each PC are computed. The identification of variables with significant loadings on a specific component is based on the rotated component matrix generated by SPSS through PCA. Using the rotated component matrix, we can identify the variables are interrelated and have strong correlations with specific PCs. The distributed assessment of each PC is computed on the basis of the component score coefficient matrix $\mathrm{A}(m * p)$ produced through PCA and generated by SPSS, and the inner logic of the computation is described as in equations (1), (2), (3), (4), and (5). The component score coefficient matrix $\mathrm{A}(m * p)$ contains $m * p$ coefficients $\alpha_{ij} (i = 1,2,\cdots m; j = 1,2,\cdots,p)$ that represent the weight or loading of the *j*th original variable for the ith extracted PC, where m is the number of extracted PCs and p is the number of surveyed items in the dataset for analysis. In this study, to compute the distributed assessment of each extracted PC, we ignore variables without significant loadings on the PC and employ only variables that load strongly on the PC. Thus, the weight $w_{ik} (i = 1,2,\cdots m; k = 1,2,\cdots,l)$ of the *k*th variable that has significant loading on the ith PC can be calculated by normalizing the corresponding coefficients of l variables, as displayed in the component score coefficient matrix, where l is the number of all variables that have significant loadings on the ith PC. The weight of the kth contributing variable for the ith PC is calculated using the following:

$$w_{ik} = \frac{\alpha_{ik}}{\sum_{k=1}^{l} \alpha_{ik}}, i = 1,2, \cdots m; k = 1,2, \cdots, l \quad (17)$$

Note that α_{ik} is always positive because we employ only variables with significant loadings on each PC to compute the distributed assessment of the PC on different grades.

We assume that the frequency distribution of the patient assessment of each surveyed item on different evaluation grades is represented as $\beta_{tj}(t = 1,2, \cdots, M; j = 1,2, \cdots, L)$, where M is the number of evaluation grades, $H_t(t = 1,2 \cdots, M)$, which are used to assess each item, and L is the number of items being assessed or surveyed. The distributed assessment of each extracted PC, Z_i, on different evaluation grades, $\beta_{Z_i,t}(i = 1,2 \cdots, m; t = 1,2, \cdots, M)$, can be computed using the following:

$$\beta_{Z_i,t} = \sum_{k=1}^{l} (w_{ik} * \beta_{tk}), t = 1,2, \cdots, M \quad (18)$$

where l is the number of all variables that have significant loadings on the ith PC, Z_i.

Finally, to aggregate the distributed assessments of extracted PCs to obtain an aggregated healthcare quality assessment result $\{(H_t, \beta_t), t = 1,2, \cdots, M\}$, the ER approach is employed on the basis of the weight of each extracted PC calculated in step five using (16) and the distributed assessment of each PC computed in step six using (17) and (18).

2.7. Combining AHP and the ER approach to assess healthcare quality

As discussed in Section 2.4, AHP is a typical method used to calculate the relative importance of criteria in a hierarchy. Therefore, AHP can be used to calculate the weights of survey items and their corresponding quality dimensions instead of using the method discussed in Section 2.6 for PC and corresponding item weight calculation in PCA.

For convenience, we used the same patient satisfaction assessment framework as determined by PCA. We consider that one PC represents one quality dimension. Therefore, the number of extracted PCs represents the number of quality dimensions that were assessed in the survey. We then used AHP to calculate the relative importance of different quality dimensions and their corresponding survey items.

We invited six domain experts to provide their judgments about the importance of quality dimensions and corresponding items in the hierarchical framework. We built pairwise comparison matrix on the basis of the respondents' responses and used the Eigenvalue method to calculate the weight of those items at different levels in the assessment framework. We then averaged the weights calculated from the experts' responses if their pairwise comparisons pass the consistency check.

After determining the weight of each quality dimension and its corresponding survey item via AHP, we employed the ER approach to aggregate the assessment of each item to obtain the overall quality assessment result.

3. Results

The characteristics of the studied survey data obtained after excluding unqualified surveys are shown in Table 1.

INSERT TABLE 1 HERE

After deleting items with a response rate lower than 90%, 19 items were retained in the dataset for analysis. The frequency of patients' evaluations of each item distributed on five-point Likert-type scales are described in Table 2.

INSERT TABLE 2 HERE

The KMO index for the studied survey dataset was 0.915 with a Bartlett's test significance of less than 0.001.

By using SPSS to perform PCA on the studied survey data, we obtained the results for the proportion of variance that is explained by each PC. We extracted seven PCs on the basis of the threshold value of 70% of the total variance that the combined PCs should account for in the dataset. The correlation between 19 items and the extracted seven PCs identified through PCA is shown in Table 3.

INSERT TABLE 3 HERE

INSERT TABLE 4 HERE

The total variance explained by the seven extracted PCs is described in Table 4. The normalized weights of the seven PCs were calculated using (16) on the basis of the eigenvalues of the seven extracted PCs. These PCs have normalized weights of w_1 = 0.561, w_2 = 0.100, w_3 = 0.084, w_4 = 0.076, w_5 = 0.066, w_6 = 0.056, and w_7 = 0.056. The rotated component matrix is shown in Table 5, where the rotated loadings of variables that strongly load on each PC are shaded gray. The component score coefficient matrix is shown in Table 6, where the coefficients of variables that strongly load on each PC are also shaded gray. These variables are used to form the linear functions used to derive the corresponding PCs.

INSERT TABLE 5 HERE

INSERT TABLE 6 HERE

On the basis of the coefficients as presented in Table 6, we calculated the weights of variables that load strongly on each PC using (17). The first PC (PC1) can be taken as an example. From Tables 5 and 6, we can identify six variables that are significantly correlated with PC1: Q5, Q6, Q8, Q9, Q10, and Q11. By normalizing their coefficients for PC1, we can obtain the corresponding weights as w_{11} = 0.309 ÷ (0.309 + 0.189 + 0.228 + 0.290 + 0.288 + 0.286) = 0.194 (Q5), w_{12} = 0.119 (Q6), w_{13} = 0.143 (Q8), w_{14}

= 0.182 (Q9), w_{15} = 0.181 (Q10), and w_{16} = 0.180 (Q11).

Next, multiplying the above calculated weights and the distributed frequency of patient evaluations on different grades as shown in Table 2 according to equation (18), we obtained the belief degrees distributed on different evaluation grades (the five-point Likert-type scales) for each PC. The distributed assessments of the seven extracted PCs are shown in Table 7.

INSERT TABLE 7 HERE

Finally, on the basis of the calculated weights and the belief degrees distributed on the five-point Likert-type scales associated with the seven extracted PCs, we employed an ER-based Intelligent Decision System (IDS) (Xu et al 2006) to model the combined healthcare quality assessment problem (Fig. 2). After aggregating the distributed assessments of the seven extracted PCs, we obtained an aggregated assessment result as shown in Fig. 3.

INSERT FIGURE 2 HERE

INSERT FIGURE 3 HERE

Alternatively, after determining the assessment framework via PCA, each PC is considered to represent one healthcare quality dimension. Thus, seven quality dimensions are assessed in the survey. We consider the following seven quality dimensions on the basis of the characteristics of items assessed in each quality dimension: 1) doctor–patient or nurse–patient communication; 2) communication about illness; 3) hospital environment; 4) admission or discharge information; 5) waiting time; 6) communication about drug or examinations; and 7) pain control or emotional support. We then employed AHP to generate the weights of the seven quality dimensions and their corresponding items.

We invited six experts to provide their preferences for the relative importance of each quality dimension and their corresponding items. In checking the consistency of the comparison matrix provided by each expert, we found that two experts' judgments are inconsistent. Therefore, we used only four experts' comparison matrix to calculate the weights of quality dimensions and their corresponding items. We used the Eigenvalue method to calculate each expert's results and averaged four experts' results to assign the final weights to each dimension and its corresponding items. The weights of the seven quality dimensions generated by AHP after averaging four experts' judgments are shown in Table 8, and the averaged weights of assessed items corresponding to each dimension are shown in Table 9.

INSERT TABLE 8 HERE

INSERT TABLE 9 HERE

Likewise, we employed IDS to aggregate the patient evaluation of each item on the basis of the weights of quality dimensions and corresponding items that we calculated through AHP. Fig. 4 shows the hierarchical assessment framework modelled by IDS in AHP method, and Fig. 5 shows the distributed assessments after aggregating all patients' evaluations based on the AHP hierarchical framework.

INSERT FIGURE 4 HERE

INSERT FIGURE 5 HERE

As the combined assessment result contains belief degrees distributed on different evaluation grades and is not straightforward enough to enable quality comparison between hospitals. Yang and Xu (2002) proposed the concept of expected utility to define a numerical value that is equivalent to the distributed assessment. For this purpose, the utilities of individual assessment grades need to be defined first. In our case, if we assign a quality score of 10 to "very satisfied," 8 to "satisfied," 6 to "fair," 4 to "dissatisfied," and 2 to "very dissatisfied," we can obtain a numerical quality score of Hospital A as $10 \times 54.47\% + 8 \times 40.11\% + 6 \times 4.43\% + 4 \times 0.76\% + 2 \times 0.22\% = 5.447 + 3.209 + 0.266 + 0.030 + 0.004 = 8.956$ through the combined method of PCA and ER. We can also obtain a quality score of 8.953 for Hospital A through the combined method of AHP and ER. If more than one hospital needs to be assessed, the numerical quality score generated for alternative hospitals can be employed to rank the healthcare quality of different hospitals.

4. Discussion and conclusions

This study proposes a new hybrid method, which combines PCA and the ER approach, for the assessment of healthcare quality based on patient experience and satisfaction surveys. In this new hybrid method, PCA helps identify the structure of the relationship between interrelated items and to derive uncorrelated PCs. The structure of the relationship among different items can be identified on the basis of the extracted PCs, and the distributed assessments of the extracted PC can be computed from corresponding variables with significant loadings. In transforming the original variables to PCs, the weights of variables are taken into account on the basis of their loadings on corresponding PCs. The ER approach is then employed to aggregate the distributed assessments of extracted PCs to obtain an overall assessment of healthcare quality. The weight of each PC is considered in aggregation and determined by the variance that the corresponding PC accounts for in the dataset.

Combining the ER approach with PCA for the aggregated assessment of healthcare quality can enhance its capability to aid MADA problems with interrelated attributes or items. Using PCA to extract PCs can help transform interrelated items into uncorrelated PCs, which can then be used as multiple attributes or criteria to be aggregated by the ER approach. In contrast to the conventional component score computation in PCA that uses all available variables in linear functions, we employ only variables that have significant loadings on the corresponding PCs to transform the original interrelated variables to PCs. The weights of variables in transformation functions are determined by their loadings on the PCs, i.e., their correlations with the corresponding PCs. This helps ensure that the distributed assessments on the extracted PCs are uncorrelated.

To compare the performance of the proposed method with that of another method, we also performed aggregated quality assessment through the combined AHP and ER approach. The quality assessment frameworks of the combined PCA and ER approach and of the combined AHP and ER approach are both derived from PCA. In the former method, the weight of each extracted PC and its corresponding items are all generated on the basis of collected data. By contrast, in the latter method, the relative importance of assessed items is calculated on the basis of the respondents' subjective judgments. These two different hybrid methods generated different aggregated distributed assessments (Fig. 3 and Fig 5) but similar overall quality scores (8.956 and 8.953).

Compared with the combined AHP and ER approach, the combined PCA and ER approach has the following advantages: it is completely based on survey data, and its result is completely objective and contains no subjective judgments . The use of AHP to calculate the weights of quality dimensions and corresponding items in the hierarchical framework has numerous disadvantages. First, an expert may have inconsistent judgments of pairwise comparison. Second, two experts may have completely different judgments for the same surveyed item set. Third, given that different experts have different opinions about healthcare quality, the weights of different dimensions and items calculated via AHP will certainly be different if different experts are surveyed. Therefore, if other experts are surveyed, we may obtain a different overall quality assessment result through the combined AHP and ER method.

In the current healthcare environment, using patient experience and satisfaction surveys to evaluate healthcare quality is necessary and integral for overall healthcare assessment. The government and general public are searching for optimal methods to assess healthcare quality from a patient's perspective, and they try to link healthcare quality assessment results to resource allocation, such as government funding support. Healthcare consumers (patients) are very interested in the ranking of the healthcare quality of different hospitals, and the hospital's quality ranking will most certainly affect patients' healthcare service choices. The new hybrid method proposed in this study provides a pragmatic and objective approach to healthcare quality assessment by aggregating patient evaluations from different dimensions or perspectives. Although only one hospital was investigated in this study, this hybrid method is suitable for assessing numerous hospitals by using the same questionnaire. Moreover, it can help rank the healthcare provided by different hospitals on the basis of various quality dimensions.

To conclude, this study proposed a novel hybrid method that combines PCA and the ER approach. The method first identifies relationships among all surveyed items from collected survey data. It then transforms original interrelated items to uncorrelated PCs. Finally, it employs the ER approach to aggregate the distributed assessments of the extracted PCs. The proposed hybrid method is objective and completely based on survey datasets. It combines the advantages of PCA and the ER approach to provide a novel and rational approach for assessing healthcare quality from the patient's perspective.

Acknowledgments

This study was supported by grants from the National Natural Science Foundation of China (grant no. 81771938, 81301296). This study was also supported by grants from Peking University (grant no. PKU2017LCX05, BMU20160592).

References

Büyüközkan G, Çifçi G. 2012. A combined fuzzy AHP and fuzzy TOPSIS based strategic analysis of electronic service quality in healthcare industry. *Expert Systems with Applications* 39: 2341-54

Büyüközkan G, Çifçi G, Güleryüz S. 2011. Strategic analysis of healthcare service quality using fuzzy AHP methodology. *Expert Systems with Applications* 38: 9407-24

Behara RS, Fisher WW, Lemmink JGAM. 2002. Modelling and evaluating service quality measurement using neural networks. *International Journal of Operations & Production Management* 22: 1162-85

Carlucci D, Renna P, Schiuma G. 2013. Evaluating service quality dimensions as antecedents to outpatient satisfaction using back propagation neural network. *Health Care Manage Science* 16: 37-44

Department of Health. 2000. The NHS Plan. London: The Stationery Office

Department of Health. 2013. Victorian health service performance monitoring framework. Victorian Government, Australia

Fragkiadakis G, Doumpos M, Zopounidis C, Germain C. 2016. Operational and economic efficiency analysis of public hospitals in Greece. *Annals of Operations Research* 247: 787-806

Goldstein E, Farquhar M, Crofton C, Darby C, Garfinkel S. 2005. Measuring hospital care from the patients' perspective: an overview of the CAHPS Hospital Survey development process. *Health Services Research* 40: 1977-95

Harris LE, Swindle RW, Mungai SM, Weinberger M, Tierney WM. 1999. Measuring patient satisfaction for quality improvement. *Medical Care* 37: 1207-13

Ishizaka A, Balkenbourg D, Kaplan T. 2010. Does AHP help us to make a choice? An experimental evaluation, . *Journal of the Operational Research Society* 62: 1801-12

Jenkinson C, Coulter A, Bruster S. 2002. The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. *International Journal for Quality in Health Care* 14: 353-8

Jenkinson C, Coulter A, Reeves R, Bruster S, Richards N. 2003. Properties of the Picker Patient Experience questionnaire in a randomized controlled trial of long versus short form survey instruments. *Journal Of Public Health Medicine* 25: 197-201

Jha AK, Orav EJ, Zheng J, Epstein AM. 2008. Patients' perception of hospital care in the United States. *The New England Journal Of Medicine* 359: 1921-31

Jolliffe IT. 2002. *Principal Component Analysis*. New York: Springer.

Keller AC, Bergman MM, Heinzmann C, Todorov A, Weber H, Heberer M. 2014. The relationship between hospital patients' ratings of quality of care and communication. *Internal Journal for*

Quality in Health Care 26: 26-33

Keller S, O'Malley AJ, Hays RD, Matthew RA, Zaslavsky AM, et al. 2005. Methods used to streamline the CAHPS Hospital Survey. *Health Services Research* 40: 2057-77

Kleefstra SM, Kool RB, Veldkamp CM, Winters-van der Meer AC, Mens MA, et al. 2010. A core questionnaire for the assessment of patient satisfaction in academic hospitals in The Netherlands: development and first results in a nationwide study. *Quality & Safety In Health Care* 19: e24

Kong GL, Xu D-L, Liu X, Yang J-B. 2009. Applying a belief rule-base inference methodology to a guideline-based clinical decision support system. *Expert Systems* 26: 391-408

Kong GL, Xu D-L, Yang J-B, Ma XM. 2015. Combined medical quality assessment using the evidential reasoning approach. *Expert Systems with Applications* 42: 5522-30

Kong GL, Xu DL, Body R, Yang JB, Mackway-Jones KRH, Carley S. 2012. A belief rule-based decision support system for clincial risk assessment of cardiac chest pain. *European Journal of Operational Research* 219: 564-73

Lyratzopoulos G, Elliott MN, Barbiere JM, Staetsky L, Paddison CA, et al. 2011. How can Health Care Organizations be Reliably Compared?: Lessons From a National Survey of Patient Experience. *Medical Care* 49: 724-33

Morgan R. 2017. An investigation of constraints upon fisheries diversification using the Analytic Hierarchy Process (AHP). *Marine Policy* 86: 24-30

Norman GR, Streiner DL. 1998. *Biostatistics: the bare essentials*. Hamilton, Ontario B.C. Decker Inc.

Panagiotis M, Kostas K, Ioannis M. 2016. Factors affecting primary health care centers' economic and production efficiency. *Annals of Operations Research* 247: 807-22

Park YS, Egilmez G, Kucukvar M. 2015. A novel life cycle-based principal component analysis framework for eco-efficiency analysis: case of the United States manufacturing and transportation nexus. *Journal of Cleaner Production* 92: 327-42

Prior D. 2006. Efficiency and total quality management in health care organizations: A dynamic frontier approach. *Annals of Operations Research* 145: 281-99

Purcărea VL, Gheorghe IR, Petrescu CM. 2013. The assessment of perceived service quality of public health care services in Romania using the SERVQUAL scale. *Procedia Economics and Finance* 6: 573-85

Rodriguez H, von Glahn T, Elliott M, Rogers W, Safran D. 2009. The Effect of Performance-Based Financial Incentives on Improving Patient Care Experiences: A Statewide Evaluation. *Journal Of General Internal Medicine* 24: 1281-88

Saaty TL. 1980. *The Analytic Hierarchy Process*. New York: McGraw-Hill.

Saaty TL. 2008. Decision making with the analytic hierarchy process *International Journal of Services Sciences* 1: 83-98

Vuković M, Gvozdenović BS, Gajić T, Stamatović Gajić B, Jakovljević M, McCormick BP. 2012. Validation of a patient satisfaction questionnaire in primary health care. *Public Health* 126: 710-18

Wang YM, Yang JB, Xu DL. 2006. Environmental impact assessment using the evidential reasoning approach. *European Journal of Operational Research* 174: 1885-913

Wong EL, Leung MC, Cheung AW, Yam CH, Yeoh EK, Griffiths S. 2011. A population-based survey using PPE-15: relationship of care aspects to patient satisfaction in Hong Kong. *International Journal for Quality in Health Care* 23: 390-6

Xu D-L. 2012. An introduction and survey of the evidential reasoning approach for multiple criteria decision analysis. *Annals of Operations Research* 195: 163-87

Xu DL, McCarthy G, Yang JB. 2006. Intelligent decision system and its application in business innovation self assessment. *Decision Support Systems* 42: 664-73

Yang JB. 2001. Rule and utility based evidential reasoning approach for multiple attribute decision analysis under uncertainty. *European Journal of Operational Research* 131: 31-61

Yang JB, Singh MG. 1994. An evidential reasoning approach for multiple-attribute decision making with uncertainty. *Systems, Man and Cybernetics, IEEE Transactions on* 24: 1-18

Yang JB, Xu DL. 2002. On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty. *IEEE Transactions on Systems Man and Cybernetics Part A-Systems and Humans* 32: 289-304

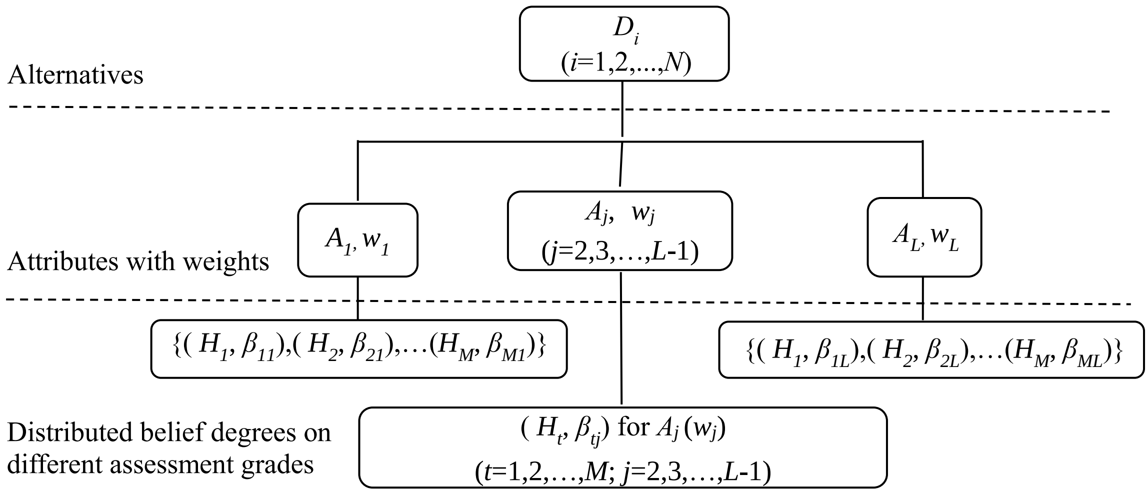

Fig. 1 A MADA problem modelled by the ER approach

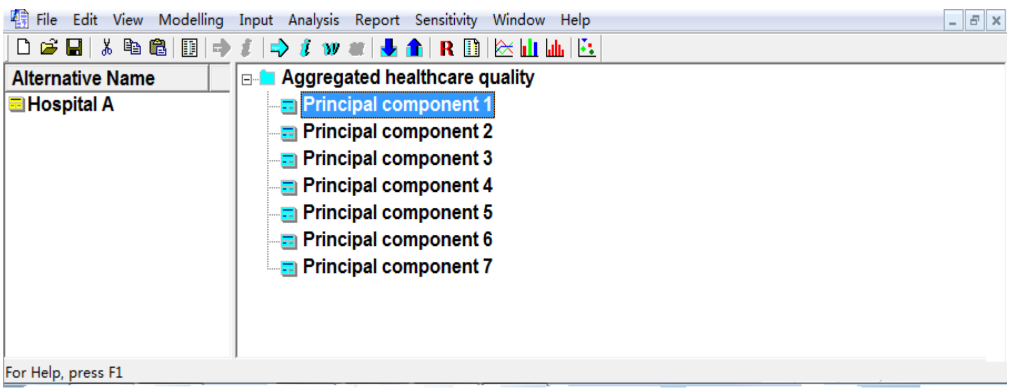

Fig. 2 The aggregated healthcare quality assessment problem modeled by IDS

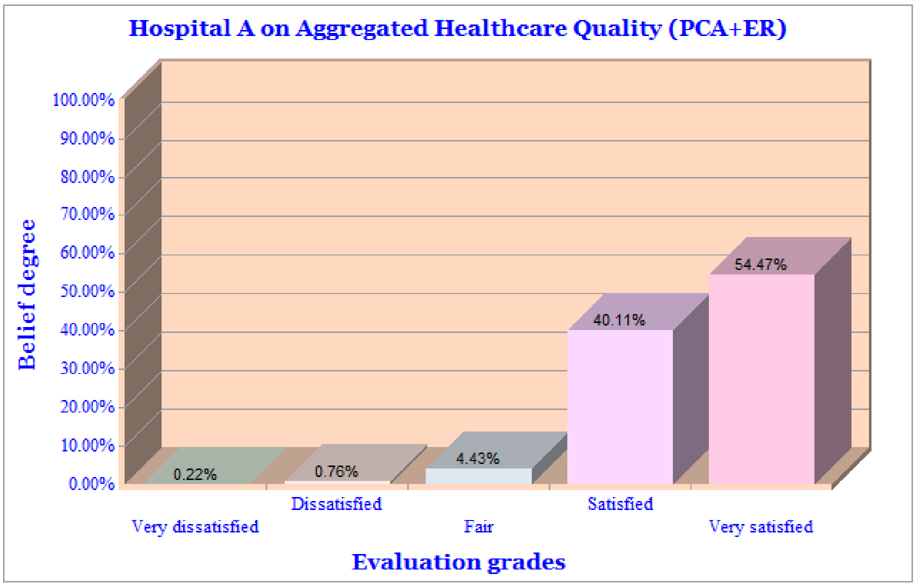

Fig. 3 The combined assessment result after aggregating assessments of the PCs

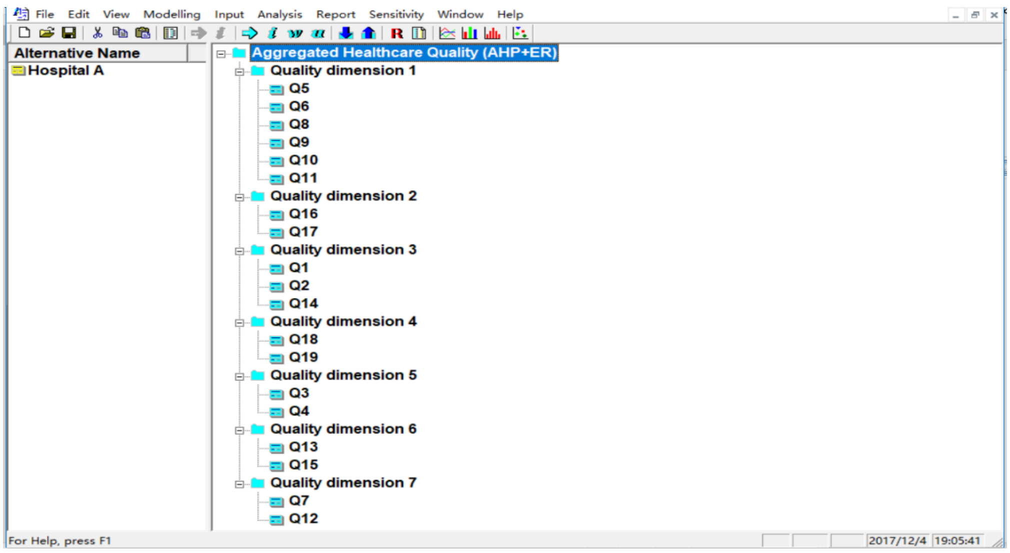

Fig. 4 The hierarchical assessment framework modeled by IDS

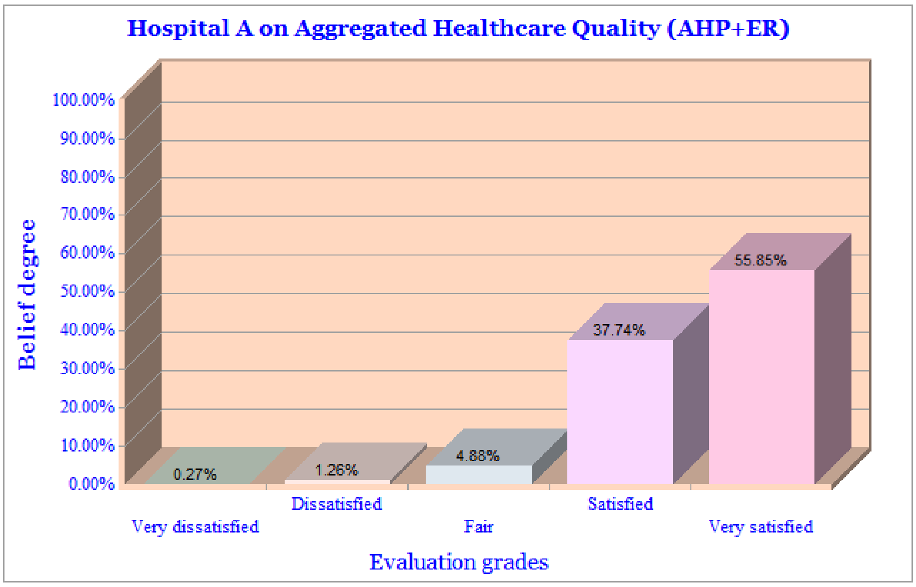

Fig. 5 The combined assessment result after aggregating evaluation on each item

Table 1 Characteristics of the studied survey data (N=192)

Variable	Subgroup	Number of patients (proportion)
Gender	Male	82(42.7%)
	Female	110(57.3%)
Age (years old)	<=44	62(32.3%)
	45-59	43(22.4%)
	60-74	57(29.7%)
	>=75	30(15.6%)
Education background	Grade school or below	27(14.1%)
	Middle school	38(19.8%)
	High school or technical school	52(27.1%)
	College or above	75(39%)
Marital status	Married	149(77.6%)
	Widowed or divorced	21(10.9%)
	Single	22(11.5%)
Health condition	Bad	10(5.2%)
	Fair	75(39.1%)
	Good	61(31.8%)
	Excellent	30(15.6%)
	Data missing	16(8.3%)
Residential Address	Beijing	139(72.4%)
	Outside Beijing	53(27.6%)

Table 2 Frequency of patients' evaluations distributed on the five-point Likert-type scales

	1- Very dissatisfied	2- Dissatisfied	3- Fair	4- Satisfied	5- Very satisfied
Q1	1.0%	1.6%	13.5%	57.3%	26.6%
Q2	1.6%	2.1%	7.3%	60.9%	28.1%
Q3	1.0%	4.2%	15.1%	51.0%	28.6%
Q4	1.0%	2.1%	14.1%	48.4%	34.4%
Q5	0	0.5%	1.6%	38.0%	59.9%
Q6	0	0.5%	2.6%	44.8%	52.1%
Q7	0.5%	3.1%	11.5%	49.0%	35.9%
Q8	0	0	2.6%	30.2%	67.2%
Q9	0	1.0%	7.3%	45.8%	45.8%
Q10	0	0.5%	3.6%	39.1%	56.8%
Q11	0	0	3.6%	43.8%	52.6%
Q12	0	1.0%	8.3%	63.5%	27.1%
Q13	0	0	9.4%	26.0%	64.6%
Q14	1.0%	1.0%	12.0%	57.3%	28.6%
Q15	0.5%	1.0%	4.7%	10.4%	83.3%
Q16	0	3.6%	5.2%	34.9%	56.3%
Q17	2.1%	3.6%	12.5%	46.4%	35.4%
Q18	1.0%	1.0%	3.1%	22.9%	71.9%
Q19	1.6%	1.0%	3.6%	13.0%	80.7%

Table 3 The correlation between the items and the extracted PCs

Component	Items measured in the questionnaire
1	Q5. Doctors treated you with respect and dignity while you were in hospital.
	Q6. Doctors gave you answers you could understand when you had important questions to ask them.
	Q8. You had trust in your doctors.
	Q9. You could get help as soon as you wanted it after you pressed the call button.
	Q10. Nurses treated you with courtesy and respect.
	Q11. Nurses explained things in a way you could understand.
2	Q16. You and your family knew about details of your condition and treatment.
	Q17. Doctors explained test results clearly to you.
3	Q1. Cleanliness of your room and bathroom.
	Q2. Convenience of using personal item lockers.
	Q14. Other hospital staff (excluding doctors and nurses) treated you with courtesy and respect.
4	Q18. Hospital staff gave you and your family enough guidance on hospital admission.
	Q19. Hospital staff gave you enough information about what symptoms or health problems to look out for after you were discharged, what activities you could and could not do, and how to take the medicine at home.
5	Q3. Time waiting to go to ward.
	Q4. Time waiting in ward for surgery to be performed.
6	Q13. Hospital staff did not bring you unexpected pain during medical examinations.
	Q15. You have been asked about your history of drug allergy and have been given enough information about the medicine, such as possible side-effects of the medicine, before giving you the medicine.
7	Q7. Doctors discussed with you when you had anxieties or fears about your condition or treatment.
	Q12. Your pain was well controlled.

Table 4 Total variance explained

Component	Initial eigenvalues		
	Total	% of variance	Cumulative %
1	7.619	40.099	40.099
2	1.361	7.165	47.263
3	1.141	6.007	53.271
4	1.030	5.420	58.691
5	0.900	4.735	63.426
6	0.763	4.014	67.440
7	0.761	4.005	71.445

Table 5 Rotated component matrix

Item	Component						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Q1	0.334	-0.012	0.689	0.210	0.140	0.072	-0.187
Q2	0.105	0.050	0.762	0.152	0.184	-0.061	0.240
Q3	0.162	0.211	0.259	0.025	0.721	-0.002	0.196
Q4	0.277	0.142	0.069	0.088	0.729	0.214	0.001
Q5	0.730	0.370	0.053	0.160	0.278	-0.009	-0.071
Q6	0.627	0.488	0.046	0.150	0.237	-0.010	0.183
Q7	0.289	0.445	0.094	0.362	0.151	-0.058	0.455
Q8	0.670	0.411	0.080	0.224	0.164	0.040	0.118
Q9	0.745	0.074	0.234	0.113	0.137	0.099	0.285
Q10	0.756	0.099	0.299	0.051	0.107	0.232	0.166
Q11	0.736	-0.006	0.292	0.129	0.108	0.224	0.197
Q12	0.357	0.132	0.125	0.104	0.143	0.207	0.709
Q13	0.437	0.307	0.047	0.221	0.004	0.487	-0.130
Q14	0.250	0.432	0.630	-0.116	-0.001	0.186	0.120
Q15	0.124	0.157	0.052	0.061	0.165	0.829	0.175
Q16	0.135	0.658	-0.042	0.181	0.316	0.224	0.202
Q17	0.206	0.705	0.218	0.155	0.105	0.184	0.005
Q18	0.191	0.213	0.179	0.727	-0.125	0.084	0.129
Q19	0.134	0.096	0.053	0.845	0.223	0.081	0.033

Table 6 Component score coefficient matrix

Item	Component						
	1	2	3	4	5	6	7
Q1	0.029	-0.134	0.448	0.111	0.050	0.013	-
Q2	-0.188	-0.060	0.521	0.044	0.058	-	0.176
Q3	-0.139	-0.036	0.081	-	0.608	-	0.070
Q4	-0.023	-0.154	-0.083	-	0.646	0.115	-
Q5	0.309	0.109	-0.142	-	0.092	-	-
Q6	0.189	0.219	-0.140	-	0.006	-	0.011
Q7	-0.068	0.196	-0.055	0.134	-	-	0.403
Q8	0.228	0.139	-0.117	0.005	-	-	-
Q9	0.290	-0.195	-0.026	-	-	-	0.170
Q10	0.288	-0.160	0.035	-	-	0.092	0.013
Q11	0.286	-0.274	0.026	-	-	0.102	0.064
Q12	-0.019	-0.139	-0.062	-	-	0.094	0.753
Q13	0.111	0.082	-0.064	0.066	-	0.394	-
Q14	-0.111	0.340	0.431	-	-	0.066	-
Q15	-0.155	-0.102	-0.033	-	0.039	0.809	0.104
Q16	-0.179	0.410	-0.112	-	0.108	0.068	0.076
Q17	-0.132	0.534	0.112	-	-	0.024	-
Q18	-0.070	0.009	0.060	0.514	-	0.008	0.044
Q19	-0.099	-0.185	-0.058	0.653	0.162	0.020	-

Table 7 Distributed assessments of the seven extracted principal components

Component	Belief degrees distributed on the five scales				
	1- Very dissatisfied	2- Dissatisfied	3- Fair	4- Satisfied	5- Very satisfied
1	0	0.45%	3.63%	40.35%	55.57%
2	1.18%	3.65%	9.33%	41.38%	44.47%
3	1.24%	1.60%	10.73%	58.65%	27.79%
4	1.33%	1.04%	3.42%	17.38%	76.83%
5	1.04%	3.09%	14.57%	49.70%	31.60%
6	0.35%	0.70%	6.22%	15.53%	77.19%
7	0.18%	1.77%	9.42%	58.46%	30.17%

Table 8 Weights of seven quality dimensions generated using AHP

Quality dimension	Weight				
	Expert1	Expert2	Expert3	Expert4	Average
1- the doctor-patient or nurse-patient communication	0.060	0.263	0.209	0.237	0.192
2- communication about illness	0.327	0.036	0.355	0.138	0.214
3- hospital environment	0.026	0.155	0.037	0.045	0.066
4- admission or discharge information	0.135	0.056	0.051	0.122	0.091
5- waiting time	0.048	0.115	0.063	0.030	0.064
6- communication about medicines or examinations	0.284	0.061	0.061	0.238	0.161
7- pain control or emotional support	0.121	0.315	0.225	0.190	0.213

Table 9 Weights of items being assessed in the survey (generated using AHP)

Dimension	Items measured in the questionnaire	Averaged weight
1	Q5. Doctors treated you with respect and dignity while you were in hospital.	0.308
	Q6. Doctors gave you answers you could understand when you had important questions to ask them.	0.221
	Q8. You had trust in your doctors.	0.220
	Q9. You could get help as soon as you wanted it after you pressed the call button.	0.108
	Q10. Nurses treated you with courtesy and respect.	0.094
	Q11. Nurses explained things in a way you could understand.	0.048
2	Q16. You and your family knew about details of your condition and treatment.	0.802
	Q17. Doctors explained test results clearly to you.	0.198
3	Q1. Cleanliness of your room and bathroom.	0.462
	Q2. Convenience of using personal item lockers.	0.260
	Q14. Other hospital staff (excluding doctors and nurses) treated you with courtesy and respect.	0.278
4	Q18. Hospital staff gave you and your family enough guidance on hospital admission.	0.500
	Q19. Hospital staff gave you enough information about what symptoms or health problems to look out for after you were discharged, what activities you could and could not do, and how to take the medicine at home.	0.500
5	Q3. Time waiting to go to ward.	0.792
	Q4. Time waiting in ward for surgery to be performed.	0.208
6	Q13. Hospital staff did not bring you unexpected pain during medical examinations.	0.333
	Q15. You have been asked about your history of drug allergy and have been given enough information about the medicine, such as possible side-effects of the medicine, before giving you the medicine.	0.667
7	Q7. Doctors discussed with you when you had anxieties or fears about your condition or treatment.	0.375
	Q12. Your pain was well controlled.	0.625