Expert Systems with Applications 42 (2015) 5522–5530

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Combined medical quality assessment using the evidential reasoning approach

Guilan Kong [a,*], Dong-Ling Xu [b], Jian-Bo Yang [b], Xiemin Ma [c]

[a] Medical Informatics Center, Peking University, Beijing, China
[b] Manchester Business School, Booth Street West, Manchester M15 6PB, UK
[c] School of Public Health, Peking University, Beijing, China

ARTICLE INFO

Article history:
Available online 21 March 2015

Keywords:
Medical quality assessment
Quality indicator
Expert judgment
Patient feedback
The evidential reasoning approach

ABSTRACT

Due to increasing demand for healthcare, medical quality has attracted significant attention in recent years. Most studies to date have tried to assess medical quality from objective quality indicators or subjective expert judgments or patient feedback perspective. In this study, the evidential reasoning approach is employed to combine objective quality indicators, subjective expert judgments and patient feedback in a multiple criteria framework to assess the quality of hospitals systematically and comprehensively. The evidential reasoning approach has the advantages of consistently handling both subjective evaluations and objective indicators under uncertainty within the same framework, and it can help to provide a robust alternative ranking. This study contributes to the literature with not only a novel medical quality assessment and aggregation framework, but also a pragmatic data transformation technique which can facilitate the combination of quantitative data and qualitative judgments using the evidential reasoning approach. A case study of three top-ranked teaching hospitals in Beijing is presented to demonstrate the framework and methodology proposed in this study.

1. Introduction

Due to increasing demand for healthcare, rising medical costs, restricted medical resources, and inevitable variations in medical practice, medical quality has attracted significant attention in recent years. Governments hope to allocate limited resources to hospitals based on medical quality (Contencin, Falcoff, & Doumenc, 2006; Normand, Wolf, & McNeil, 2008; Rees & Dineschandra, 2005; Ritchey et al., 2012). The public craves high-quality healthcare services, and people choose appropriate hospitals on the basis of medical quality information they can collect about target hospitals (Dijs-Elsinga et al., 2010; Glazer, McGuire, Cao, & Zaslavsky, 2008; Marang-van de Mheen et al., 2011). Meanwhile, hospital managers seek to improve medical quality, as quality is the key factor for attracting public or private funding and healthcare service consumers (Campbell, Roland, & Buetow, 2000; Carlucci, Renna, & Schiuma, 2013; Glazer et al., 2008; Normand et al., 2008). In developed countries such as the US, healthcare researchers have conducted systematical medical quality research since the 1960s (Donabedian, 1966; Donabedian, 1968; Feinstein, 2002; Mcqueen, Mittman, & Demakis, 2004). However, research on medical quality in developing countries lags far behind. In China, due to rapid economic growth in the past several decades, medical quality has been attracting increasing attention from both the government and the public. In 2009, the Chinese government launched a new wave of healthcare reform. The reform was intended to reduce healthcare costs, and improve healthcare quality and patient safety. To achieve these goals, the current healthcare strategy in China links the medical quality of hospitals with the allocation of healthcare resources such as government funding. Therefore, how to assess medical quality objectively and comprehensively so as to achieve a convincing quality ranking of hospitals has become a hot current research topic in China.

In the literature, a globally accepted medical quality framework was proposed by Donabedian (1966), Donabedian (1968), who suggested that medical quality could be assessed from the aspects of medical structure (MS), medical processes (MP), and medical outcomes (MO). Donabedian's medical quality model has become a practical and standard framework for medical quality researchers since it was first proposed. Although arguments are still being made about whether it is better to assess medical quality from a process perspective or from an outcome perspective (Feinstein, 2002; Ploeg, Flu, Lardenoye, Hamming, & Breslau, 2010), objective indicator methods, or subjective expert judgments, or patient

* Corresponding author. Tel.: +86 10 82805902.
E-mail addresses: Guilan.kong@hsc.pku.edu.cn (G. Kong), ling.xu@mbs.ac.uk (D.-L. Xu), jian-bo.yang@manchester.ac.uk (J.-B. Yang), ma_xm@126.com (X. Ma).

http://dx.doi.org/10.1016/j.eswa.2015.03.009
0957-4174/

feedback have been employed to assess medical quality under Donabedian's framework (Feinstein, 2002; Kerr et al., 2007; Untachai, 2013). We briefly discuss the three types of assessment methods as follows.

In objective indicator methods, indicators are derived from electronic medical records (EMRs), electronic health records (EHRs) or administrative data sets. Normand et al. (2008), Benin et al. (2005), and Cebul, Love, Jain, and Hebert (2011) conducted medical quality assessment based on EMRs or EHRs, while Bellows and Halpin (2008) extracted quality indicators from administrative data sets. Some studies (e.g. MacLean et al., 2006) show that there is no obvious bias in using the two different data sets if similar indicators are employed, and clinical data performs better if the totality of care that can be measured by each data source is measured. From an outcome aspect, indicators used for quality assessment include hospital readmission rate (Halfon et al., 2006; Weissman et al., 1999; Wray, Peterson, Souchek, Ashton, & Hollingsworth, 1997), hospital mortality rate (Baker et al., 2002; Bottle & Aylin, 2008; Chae, Kim, Tark, Park, & Ho, 2003; Glance, Dick, Mukamel, Li, & Osler, 2010; Hofer & Hayward, 1996; Kipnis, Escobar, & Draper, 2010; Rosenthal, Shah, Way, & Harper, 1998; Thomas & Hofer, 1999), and some other negative indexes (Heineken, Charles, Stimson, Wenell, & Stimson, 1985). From the aspects of structure and process, indicators used for assessment can be adherence to practice guidelines (Ashton et al., 1994).

Expert judgments are frequently used in practice audits, peer reviews, or practice visits for assessing medical quality (Pearson et al., 2000). As no EMRs or administrative data can reflect the entire medical process or medical outcome, subjective expert judgments are complementary to objective indicators in medical quality assessment. Research shows that there is moderate to high agreement between these two different types of assessment methods (Kerr et al., 2007).

In addition to the aforementioned two assessment methods, patient feedback on health care is a measure of patient perception of the quality of care, and therefore it is considered to be an important outcome of health care and essential element of quality assessment (Chang & Chang, 2013; Donabedian, 1966; Farley et al., 2014; Untachai, 2013; Vasudevan, Arachchi, & van Langenberg, 2013). In developed countries such as Britain, assessing patients' evaluations of health care has been a requirement for some general practitioners (Pouwer & Snoek, 2002).

Due to the fact that single indicators derived from objective data, subjective expert judgments or patient feedback can only measure medical quality from limited perspectives, they may not be able to reflect the quality of medical care of one hospital as a whole (Rosenthal et al., 1998). In the literature, various aggregation methods have been employed to combine multiple quality indicators. Reeves et al. (2007) and Normand et al. (2008) used simple average sum methods to aggregate multiple quality clinical indicators. Goodson and Jang (2008) employed the Bayesian network to combine multiple objective quality factors. Büyüközkan, Çifçi, and Güleryüz (2011) extended traditional analytic hierarchy process (AHP) methodology to a fuzzy AHP to combine subjective and vague expert judgments about multiple quality factors. Büyüközkan and Çifçi (2012) combined a fuzzy AHP and a fuzzy technique for order performance by similarity to ideal solution (TOPSIS) to aggregate patient feedback on multiple quality factors.

However, each aggregation method used in the literature has its merits and limitations. More specifically, simple average sum methods are indeed simple to implement, but the combinatorial contribution or orthogonal sum of multiple factors cannot be reflected in the result. Bayesian network has the advantages of using powerful algorithms for probabilistic inference. However, the complexity of a Bayesian network increases exponentially with the increase of parameters used in the network and also Bayesian inference depends on prior distributions, the credibility of which in turn relies heavily on sampling method used for data collection. These requirements make it difficult for researchers to generate necessary parameter values or probabilities to conduct robust analysis in real life applications. Fuzzy AHP extends the traditional AHP method in dealing with vague subjective judgments about multiple criteria and has the advantages of converting subjective judgments to numerical values, but the problem of rank reversal exists in the method, which means that the ranking of alternatives may change when new alternatives are added. Similar to fuzzy AHP method, fuzzy TOPSIS method has the advantages of handling fuzzy judgments about multiple criteria, but it has the problem of rank reversal as well. Moreover, the aforementioned aggregation methods are for combining either multiple objective indicators or multiple subjective evaluations, and few studies in the literature have dealt with aggregating a mixture of objective indicators, subjective expert judgments and patient feedback to produce a more comprehensive and informative quality assessment result.

In this study, we propose to aggregate objective quality indicators, subjective expert judgments and patient feedback to assess medical quality (Kong, Ma, Zhao, & Zhang, 2013), and the work is conducted under Donabedian's medical quality framework. The indicators that we use in this study are from a MO perspective including *inpatient mortality rate* (IMR), *readmission rate* (RR), and *adverse event rate* (AER). The data sources for deriving objective indicators are the inpatient medical record summaries (IMRSs) from January 2006 to December 2010 of three top-ranked teaching hospitals in Beijing. We invited 10 area experts from hospitals and universities to provide anonymous judgments about the medical quality of the studied hospitals from the medical facilities (MF), medical staff (MSf), MP, and MO perspectives. Note here that we use MF and MSf as substitutes for MS. Furthermore, we surveyed a random sample of patients who were believed to have visited the studied hospitals during the study period using questionnaires via WeChat (http://www.wechat.com/en/). The analytic approach we employed to aggregate objective indicators, subjective judgments and patient feedback is the evidential reasoning (ER) approach (Yang & Singh, 1994; Yang & Xu, 2002). The ER approach, which provides a modeling framework and analysis method for handling both qualitative and quantitative attributes under uncertainty, has the advantages of dealing with both subjective evaluations and objective indicators under uncertainty such as vagueness or incompleteness, and can overcome the shortcomings of those aggregation methods as discussed above. In this paper, the ER approach is applied to assess the overall quality of medical care of hospitals based on both subjective evaluations and objective quantitative indicators for the first time, and a pragmatic method for transforming numerical indicators to qualitative assessment grades with a belief structure is proposed.

In the remainder of the paper, the Methods section introduces the source data, the indicators, the questionnaires that we used for acquiring expert judgments and patient feedback, the ER approach, and detailed data aggregation procedures. The Results section presents the medical quality assessment results based on IMRSs, expert judgments and patient feedback. Specifically, quality trends from 2006 to 2010 and quality ranking per year of the studied hospitals are provided. Finally, the Discussion and conclusions section summarizes the contributions and limitations of this study, and suggests future research directions.

2. Methods

2.1. Data

Because of legal and ethical concerns, it is difficult to obtain EMRs or EHRs data in China. For administration purposes, IMRSs

are the only data from hospitals that are required by the government and submitted electronically and regularly to health bureaus at different levels in China. The government has developed a standardized and computerized form to collect IMRSs data from different hospitals, and the yearly IMRSs data of each hospital is consistent in the country's record. Therefore, IMRSs data is an objective and reliable source for computation of outcome indicators, and we chose IMRSs as the data source in this study. We obtained the IMRSs from the Beijing Municipal Bureau of Health (BMBH), and the data were submitted by three top-ranked teaching hospitals in Beijing from 2006 to 2010. IMRSs contain summarized inpatient data including patient demographics, admission date, discharge date, state of illness at admission, admission diagnosis, discharge diagnosis, state of hospital infection, operations, state of illness on discharge, and so on.

On request from the studied hospitals, we use Hospital A, Hospital B, and Hospital C to anonymously represent the three studied hospitals in the paper.

2.2. Quality indicators

Traditionally, quality indicators including overall hospital mortality rate (Kipnis et al., 2010), readmission rate (Halfon et al., 2006; Weissman et al., 1999), and adverse events rate (Mull et al., 2014) regardless of differences in diseases and patients were used as measures to assess the overall quality of medical care of hospitals. Furthermore, for more accurate hospital quality assessment, main diseases or operations for which the outcomes are most likely to be causally related to antecedent medical care were selected to calculate indicators separately or aggregately for quality comparison (Baker et al., 2002; Glance et al., 2010; Wray et al., 1997). In this study, indicators that can be derived from IMRSs include IMR, RR, and AER. As most hospitalizations are caused by a few major types of diseases and operations, we selected seven main types of diseases and seven main types of operations in China, for which the outcomes are causally affected by the quality of medical care provided by hospitals, for deriving IMR and RR. We also selected four main adverse events for deriving AER. The selected main types of disease include acute myocardial infarction, cerebral hemorrhage, cerebral infarction, bacterial pneumonia, chronic obstructive pulmonary disease, congenital heart disease, and acute lymphatic leukemia. The selected main types of operation include hip replacement and knee arthroplasty, coronary artery bypass grafting, percutaneous coronary intervention, craniocerebral operations, uterine-incision delivery, exploratory laparotomy, and cardiac valve replacement. The selected main adverse events for deriving AER indicator include inpatient pressure ulcer, post-operation pulmonary infection, post-operation complication, and post-operation pulmonary embolism. Since the purpose of this study is to assess the overall quality of care of the studied hospitals, individual indicators need to be aggregated to represent the quality of each studied hospital. We used International Classification of Diseases (ICD-10) codes to identify the clinical cases of different diseases and operations from IMRSs. Note that the indicators used in this study are for general hospitals, and they may not suit the specialized hospitals, because some of the main diseases or operations selected for indicator calculation for general hospitals may not be treated or performed in specialized hospitals.

2.3. Questionnaires for acquiring expert judgments and patient feedback

In addition to using explicit indicators computed from IMRSs to assess medical quality, we also invited area experts to provide judgments about medical quality in the studied hospitals. In the designed questionnaire for acquiring expert views, items to be assessed in the studied hospitals include MF, MSf, MP, and MO. Experts can provide their judgments about the assessed items by selecting an option from 5-level grades: *excellent*, *good*, *average*, *poor*, and *worst*. The questionnaire is shown in Table 1.

Furthermore, we also used questionnaires to survey patients' feedback on the quality of health care of the studied hospitals during the studied period from MF, MSf, MP, and MO perspectives. The platform provided by WeChat was used for questionnaire delivery and collection. The questionnaire for acquiring patient feedback on the quality of the studied hospitals is similar to the questionnaire for acquiring expert judgments. The patient feedback questionnaire is shown in Table 2.

2.4. The ER approach

After derivation of objective indicators and acquisition of expert judgments and patient feedback, we employed the ER approach to aggregate data. The ER approach is a generic evidence-based multiple attribute decision analysis (MADA) approach for dealing with problems having both qualitative and quantitative attributes under uncertainty (Yang & Singh, 1994; Yang & Xu, 2002). It has been widely used in various areas such as motorcycle evaluation (Yang & Singh, 1994), bridge condition assessment (Wang & Elhag, 2008), nuclear waste repository assessment (Xu, 2009), environmental impact assessment (Wang, Yang, & Xu, 2006), weapon system capability assessment (Jiang, Li, Zhou, Xu, & Chen, 2011), and clinical risk assessment (Kong et al., 2012).

Differing from conventional MADA approaches, a belief structure was introduced in the ER approach. Assuming there are N alternatives $D(D_1, D_2, \ldots, D_N)$ that need to be assessed or ranked based on L attributes or factors $A(A_1, A_2, \ldots, A_L)$, the ℓth attribute $A_l (l = 1, 2, \ldots, L)$ can be either qualitative or quantitative, and each attribute A_l can be assessed through a set of assessment grades $H(H_1, H_2, \ldots, H_M)$ which are assumed to be collectively exhaustive and mutually exclusive. Considering the fact that the attributes $A(A_1, A_2, \ldots, A_L)$ may be of different importance, attribute weight $\omega_l (l = 1, 2, \ldots, L)$ can be employed to denote such different importance, and they should meet the condition of $\omega_l \geqslant 0$ and $\sum_{l=1}^{L} \omega_l = 1$. Assuming $\beta_{ml} (m = 1, 2, \ldots, M; l = 1, 2, \ldots, L)$ is the degree of belief in the mth assessment grade H_m on assessment of the lth attribute A_l, it can either be subjective if it quantifies a "personal belief" or objective if it is a computed probability on the basis of recorded data. A belief decision matrix (Yang & Xu, 2002) can be used to describe the performance assessment problem modeled by the ER approach as shown in Table 3.

Table 1
The questionnaire for acquiring expert judgments.

Questions about medical quality	Assessment grades				
Q^1: What do you think about the quality of medical facilities in the hospital?	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
Q^2: What do you think about the quality of medical staff in the hospital?	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
Q^3: What do you think about the quality of medical processes in the hospital?	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
Q^4: What do you think about the medical outcomes in the hospital?	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst

Table 2
The questionnaire for acquiring patient feedback.

Questions		Answers				
Q^1: Did you visit Hospital A during 2006–2010?		☐ Yes	☐ No			
(If yes to Q^1) your impression of Hospital A	Medical facilities	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
	Medical staff	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
	Medical processes	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
	Medical outcomes	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ worst
Q^2: Did you visit Hospital B during 2006–2010?		☐ Yes	☐ No			
(If yes to Q^2) Your impression of Hospital B	Medical facilities	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
	Medical staff	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
	Medical processes	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
	Medical outcomes	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ worst
Q^3: Did you visit Hospital C during 2006–2010?		☐ Yes	☐ No			
(If yes to Q^3) Your impression of Hospital C	Medical facilities	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
	Medical staff	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
	Medical processes	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst
	Medical outcomes	☐ Excellent	☐ Good	☐ Average	☐ Poor	☐ Worst

Based on the belief decision matrix, the ER algorithm can be employed to aggregate the distributed assessments of all attributes and generate a combined assessment of an alternative. The recursive ER algorithm (Yang & Singh, 1994; Yang & Xu, 2002) is briefly introduced as follows.

First, transform the degrees of belief β_{ml} $(m = 1, 2, \ldots, M, l = 1, 2, \ldots, L)$ into basic probability mass by combining the relative weights and the degrees of belief using the following equations:

$$m_{m,l} = \omega_l \beta_{ml} \tag{1}$$

$$m_{H,l} = 1 - \sum_{m=1}^{M} m_{m,l} = 1 - \omega_l \sum_{m=1}^{M} \beta_{ml} \tag{2}$$

$$\overline{m}_{H,l} = 1 - \omega_l \tag{3}$$

$$\tilde{m}_{H,l} = \omega_l \left(1 - \sum_{m=1}^{M} \beta_{ml}\right) \tag{4}$$

where $m_{H,l} = \overline{m}_{H,l} + \tilde{m}_{H,l}$ for all $l = 1, \ldots, L$ and $\sum_{l=1}^{L} \omega_l = 1$. $m_{m,l}$ represents the basic probability mass of A_l being assessed to the assessment grade H_m. Note that the probability mass assigned to the grade set H, which is unassigned to any individual attribute, is split into two parts: one caused by the relative importance of the lth attribute A_l (or $\bar{m}_{H,l}$) and the other by the incompleteness of the lth attribute A_l (or $\tilde{m}_{H,l}$). $\bar{m}_{H,l}$ represents how much the other attributes can contribute to assessing an alternative and it is the proportion of beliefs that remain to be assigned depending upon how other attributes are assessed. In essence, $\bar{m}_{H,l}$ provides a scope for conflict resolution in the presence of conflicting evidence. $\tilde{m}_{H,l}$ will be zero if there is no ignorance in the assessment.

Then, all the L attributes are aggregated to generate the combined degree of belief in each possible grade H_m. Suppose $m_{m,I(l)}$ is the combined degree of belief in H_m by aggregating the first l attributes $(A_1, A_2, \ldots, A_l)$, and $m_{H,I(l)}$ is the remaining degree of belief unassigned to any grade. Let $m_{m,I(1)} = m_{m,1}$ and $m_{H,I(1)} = m_{H,1}$. Then, the overall combined degree of belief β_m in H_m is generated as follows:

$$\{H_m\} : m_{m,I(l+1)} = K_{I(l+1)}[m_{m,I(l)} m_{m,l+1} + m_{m,I(l)} \times m_{H,l+1} + m_{H,I(l)} m_{m,l+1}], l = 1, 2, \ldots, L-1 \quad m_{H,I(l)} = \bar{m}_{H,I(l)} + \tilde{m}_{H,I(l)}, l = 1, 2, \ldots, L \tag{5}$$

$$\{H\} : \tilde{m}_{H,I(l+1)} = K_{I(l+1)}[\tilde{m}_{H,I(l)} \tilde{m}_{H,l+1} + \tilde{m}_{H,I(l)} \times \tilde{m}_{H,l+1} + \bar{m}_{H,I(l)} \tilde{m}_{H,l+1}], \quad l = 1, 2, \ldots, L-1 \tag{6}$$

$$\{H\} : \bar{m}_{H,I(l+1)} = K_{I(l+1)}[\bar{m}_{H,I(l)} \bar{m}_{H,l+1}], \quad l = 1, 2, \ldots, L-1 \tag{7}$$

$$K_{I(l+1)} = \left[1 - \sum_{m=1}^{M} \sum_{\substack{t=1 \\ t \neq m}}^{M} m_{m,I(l)} m_{t,l+1}\right]^{-1}, \quad l = 1, 2, \ldots, L-1 \tag{8}$$

$$\{H_m\} : \beta_m = \frac{m_{m,I(L)}}{1 - \bar{m}_{H,I(L)}}, m = 1, 2, \cdots, M \tag{9}$$

$$\{H\} : \beta_H = \frac{\tilde{m}_{H,I(L)}}{1 - \bar{m}_{H,I(L)}}. \tag{10}$$

β_H represents the remaining belief degrees unassigned to any H_m. It has been proven that $\sum_{m=1}^{M} \beta_m + \beta_H = 1$ (Yang & Xu, 2002). The aggregated assessment can be denoted by $S(A_l(l = 1, 2, \ldots, L)) = \{(H_m, \beta_l) m = 1, 2, \ldots, M, l = 1, 2, \ldots, L\}$.

Based on the distributed assessment of each attribute, we can identify weak areas together with strengths for each alternative. Furthermore, to rank alternatives on one attribute or all attributes, a single score to represent the performance of each alternative is necessary. Distributed assessment results, as discussed above, may not be directly used for ranking. Yang and Xu (2002) proposed employing the concept of expected utility to generate a numerical value from each distributed assessment to rank alternatives. For example, if the overall medical quality of one hospital is assessed as {(*excellent*, 0.85), (*good*, 0.00), (*average*, 0.15), (*poor*, 0.00), (*worst*, 0.00)}, and we assign a utility of 100 to *excellent* quality, 80 to *good* quality, 60 to *average* quality, 40 to *poor* quality, and 20 to *worst* quality, then we can obtain a combined quality score of the hospital as follows: $0.85 * 100 + 0.00 * 80 + 0.15 * 60 + 0.00 * 40 +$

Table 3
A belief decision matrix in the ER approach.

Attribute evaluation grades	Belief degrees					
	$(\omega_1)A_1$	$(\omega_2)A_2$	$\ldots$	$(\omega_l)A_l$	$\ldots$	$(\omega_L)A_L$
H_1	β_{11}	β_{12}	$\ldots$	β_{1l}	$\ldots$	β_{1L}
H_2	β_{21}	β_{22}	$\ldots$	β_{2l}	$\ldots$	β_{2L}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$
H_m	β_{m1}	β_{m2}	$\ldots$	β_{ml}	$\ldots$	β_{mL}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$
H_M	β_{M1}	β_{M2}	$\ldots$	β_{Ml}	$\ldots$	β_{ML}

Table 4
Expert judgments about medical quality of Hospitals A, B, and C (2006–2010).

Hospitals	Quality items	Expert judgments
Hospital A	MF	{(excellent, 100%), (good, 0), (average, 0), (poor, 0), (worst, 0)}
	MSf	{(excellent, 0), (good, 50%), (average, 50%), (poor, 0), (worst, 0)}
	MP	{(excellent, 0), (good, 60%), (average, 40%), (poor, 0), (worst, 0)}
	MO	{(excellent, 0), (good, 90%), (average, 0), (poor, 10%), (worst, 0)}
Hospital B	MF	{(excellent, 0), (good, 70%), (average, 30%), (poor, 0), (worst, 0)}
	MSf	{(excellent, 0), (good, 80%), (average, 20%), (poor, 0), (worst, 0)}
	MP	{(excellent, 0), (good, 50%), (average, 50%), (poor, 0), (worst, 0)}
	MO	{(excellent, 0), (good, 60%), (average, 40%), (poor, 0), (worst, 0)}
Hospital C	MF	{(excellent, 50%), (good, 30%), (average, 20%), (poor, 0), (worst, 0)}
	MSf	{(excellent, 0), (good, 60%), (average, 40%), (poor, 0), (worst, 0)}
	MP	{(excellent, 0), (good, 50%), (average, 40%), (poor, 10%), (worst, 0)}
	MO	{(excellent, 0), (good, 40%), (average, 50%), (poor, 10%), (worst, 0)}

Table 5
Patient feedback on medical quality of Hospitals A, B, and C (2006–2010).

Hospitals	Quality Items	Patient feedback
Hospital A	MF	{(excellent, 19.75%), (good, 37.04%), (average, 39.51%), (poor, 2.47%), (worst, 1.23%)}
	MSf	{(excellent, 18.52%), (good, 39.51%), (average, 35.80%), (poor, 3.70%), (worst, 2.47%)}
	MP	{(excellent, 12.35%), (good, 33.33%), (average, 49.38%), (poor, 4.94%), (worst, 0)}
	MO	{(excellent, 14.81%), (good, 48.15%), (average, 30.86%), (poor, 6.17%), (worst, 0)}
Hospital B	MF	{(excellent, 19.72%), (good, 53.52%), (average, 23.94%), (poor, 2.82%), (worst, 0)}
	MSf	{(excellent, 19.72%), (good, 49.30%), (average, 28.17%), (poor, 2.81%), (worst, 0)}
	MP	{(excellent, 19.72%), (good, 45.07%), (average, 35.21%), (poor, 0), (worst, 0)}
	MO	{(excellent, 21.13%), (good, 47.89%), (average, 28.17%), (poor, 2.81%), (worst, 0)}
Hospital C	MF	{(excellent, 23.68%), (good, 50.00%), (average, 22.37%), (poor, 2.63%), (worst, 1.32%)}
	MSf	{(excellent, 26.32%), (good, 47.37%), (average, 21.05%), (poor, 5.26%), (worst, 0)}
	MP	{(excellent, 19.74%), (good, 43.42%), (average, 31.58%), (poor, 3.95%), (worst, 1.31%)}
	MO	{(excellent, 27.63%), (good, 46.05%), (average, 23.68%), (poor, 1.32%), (worst, 1.32%)}

$0.00 * 0 = 94$. The quality score of 94 can be used to rank the medical quality of the hospital.

Generally speaking, there are three advantages in employing the ER approach for performance assessment. Firstly, it provides a novel belief framework to model and synthesize both objective and subjective information under uncertainty. Secondly, the ER approach can produce distributed assessments on each alternative together with its contributing attributes or criteria, and thus can help to find the weaknesses and strengths of each alternative. Thirdly, the numerical scores generated from the ER approach based on the distributed assessments can help to rank alternatives. Moreover, the ER approach requires only judgmental independence among contributing attributes, which is relatively easy to check (Xu, McCarthy, & Yang, 2006).

2.5. Analysis

The ER approach as introduced above fits just right for helping solve the medical quality assessment problem. Based on the collected IMRSs from 2006 to 2010, the acquired expert judgments and patient feedback, we conducted data aggregation via the ER approach step by step as follows.

Step 1: Extract quantitative indicators from IMRSs, and transform numerical indicator values to assessment grades with a belief structure.

First of all, we extracted indicators including IMR, RR, and AER from IMRSs for each hospital from 2006 to 2010. Regarding IMR and RR, we selected seven main types of disease and seven main types of operation as discussed before to derive indicators. For simplicity, we used the same set of assessment grades *excellent*, *good*, *average*, *poor*, and *worst* to assess both subjective evaluations and objective indicators. To transform numerical indicator values to distributed assessments with belief degrees equivalently, we need to calculate the benchmark values of *excellent*, *good*, *average*, *poor*, and *worst* grades for each indicator. For this purpose, we proposed a pragmatic method for transforming numerical values to assessment grades with a belief structure. First of all, we collected IMRSs from a sample of 30 hospitals of similar sizes from BMBH. Based on the sample data, we computed the minimum (denoted by a), the 25th percentile (b), the 50th percentile (c), the 75th percentile (d) and the maximum (e) for each indicator. As the indicators of IMR, RR, and AER all reflect negative aspects of medical quality, smaller indicator values mean better medical quality. We then set the computed a, b, c, d, and e for each indicator as benchmark values at *excellent*, *good*, *average*, *poor*, and *worst* grades respectively. Based on the benchmark values, we transformed the numerical value of each indicator to a distributed assessment using belief degrees (Yang, 2001). If we use z to represent the numerical value of an indicator and α, β, γ, δ, and θ to represent the degrees of belief in *excellent*, *good*, *average*, *poor* and *worst* grades respectively after transforming numerical value z to assessment grades, the detailed data transformation method can be described as follows, which is based on the principle of equivalence between the numerical value z and the transformed distribution (Yang, 2001).

(1) If z is less than the benchmark value at *excellent* grade ($z < a$), then the indicator can be definitely assessed as *excellent* grade, and z can be transformed to assessment with belief degree of 1 ($\alpha = 1$) associated with *excellent*. Belief degrees assigned to other grades: *good*, *average*, *poor* and *worst* are all set to be 0 ($\beta = 0$, $\gamma = 0$, $\delta = 0$, $\theta = 0$).

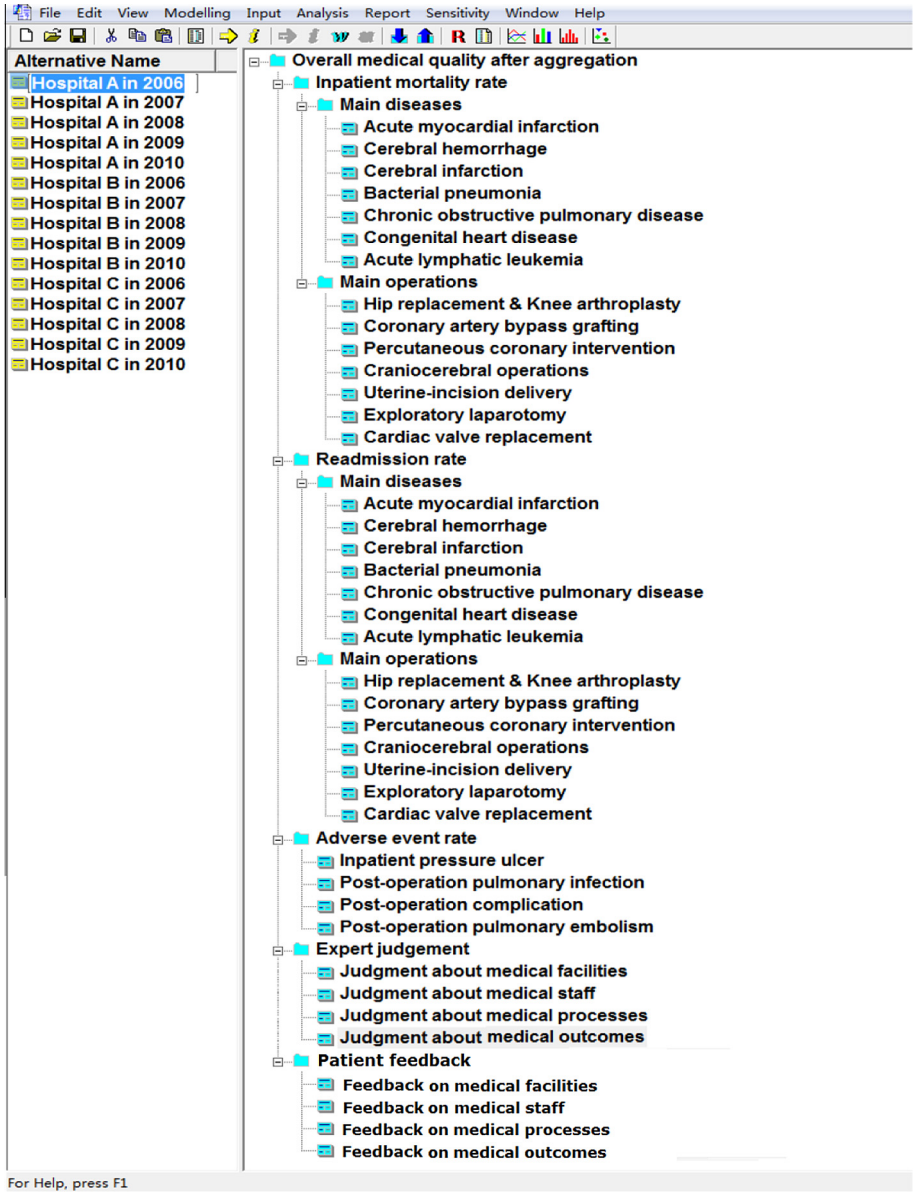

Fig. 1. The medical quality assessment framework modeled by IDS.

(2) If z is equal or greater than the benchmark value for the *excellent* grade and less than the benchmark value for the *good* grade ($a \leqslant z < b$), then z can be transformed to assessment with belief degrees ($\alpha = (b - z)/(b - a) \times 100\%$) and ($\beta = 1 - \alpha$) associated with the *excellent* and *good* grades respectively. Belief degrees assigned to the other grades: *average*, *poor* and *worst* are all set to be 0 ($\gamma = 0$, $\delta = 0$, $\theta = 0$).

(3) If $b \leqslant z < c$, then the transformed belief degrees associated with the *good* and *average* grades are ($\beta = (c - z)/(c - b) \times 100\%$) and ($\gamma = 1 - \beta$) respectively. Belief degrees assigned to the other grades: *excellent*, *poor* and *worst* are all set to be 0 ($\alpha = 0$, $\delta = 0$, $\theta = 0$).

(4) If $c \leqslant z < d$, then the transformed belief degrees associated with the *average* and *poor* grades are ($\gamma = (d - z)/(d - c) \times 100\%$) and ($\delta = 1 - \gamma$) respectively. Belief degrees assigned to the other grades: *excellent*, *good* and *worst* are all set to be 0 ($\alpha = 0$, $\beta = 0$, $\theta = 0$).

(5) If $d \leqslant z < e$, then the transformed belief degrees associated with the *poor* and *worst* grades are ($\delta = (e - z)/(e - d) \times 100\%$) and ($\theta = 1 - \delta$) respectively. Belief degrees assigned to other grades: *excellent*, *good* and *average* are all set to be 0 ($\alpha = 0$, $\beta = 0$, $\gamma = 0$).

(6) If $z \geqslant e$, then the transformed belief degree associated with the *worst* grade is 1 ($\theta = 1$). Belief degrees assigned to the other grades: *excellent*, *good*, *average* and *poor* are all set to be 0 ($\alpha = 0$, $\beta = 0$, $\gamma = 0$, $\delta = 0$).

Step 2: Acquire expert judgments.

Due to the fact that it is difficult for an expert to retrospectively judge the yearly medical quality of one hospital in the past few years, we made three assumptions about expert judgments. Firstly, we assumed that an expert's judgment about the yearly medical quality of each studied hospital remains the same during the period between 2006 and 2010. Secondly, we assumed that

all experts are of equal importance. Finally, we assumed that the 10 experts provide their judgments independently. The overall judgments of the 10 experts about quality aspects in the studied hospitals are summarized in Table 4.

In Table 4, the numerical value associated with each assessment grade represents the percentage of experts who provided the corresponding judgment. For example, {(*excellent*,0), (*good*,50%), (*average*,50%), (*poor*,0), (*worst*,0)} corresponding to "MSf" of Hospital A means 50% of the experts judge the medical staff in Hospital A as "*good*," and the other 50% experts judge it as "*average*".

Step 3: Acquire patient feedback.

Ideally, a patient's feedback on the quality of medical care of one hospital should be surveyed within half a year after the patient leaves the hospital. Due to the lack of timely patient survey during the studied period for the hospitals, we conducted a separate patient survey and assumed that each patient's feedback for one hospital reflects his/her perception of annual quality status of the hospital and a patient's yearly perception remains the same during the studied period. In total, a random sample of nearly 300 adults was surveyed, and 222 responded to the survey. Among the 222 adults, 81 responded to Hospital A, 71 to Hospital B, and 76 to Hospital C. The patient feedback on the quality of the studied hospitals is summarized in Table 5.

Step 4: Aggregate objective indicators, subjective expert judgments and patient feedback using the ER approach.

Before data aggregation, we assigned equal weights to all objective indicators, expert judgments, and patient feedback. We used MATLAB to develop a computerized program to calculate and aggregate all indicators and subjective evaluations automatically. For illustration, the ER-based Intelligent Decision System (IDS) (Xu et al., 2006) was used to show the hierarchical structure of the medical quality assessment framework based on objective indicators, expert judgments and patient feedback. Fig. 1 shows the hierarchical structure of the quality assessment framework in IDS.

After combining all objective indicators and subjective evaluations, we got distributed assessments about the yearly medical quality of the studied hospitals from 2006 to 2010 (see Table 6).

Step 5: Rank medical quality of the three hospitals.

To rank medical quality of different hospitals, it is desirable to generate numerical values that are equivalent, in terms of expected utility, to the distributed assessments. For this purpose, the utilities of individual assessment grades need to be defined first (Yang & Xu, 2002). In our study, we used quality scores to define such utility values of different assessment grades. More specifically, we assigned a quality score of 100 to *excellent* quality, 80 to *good*, 60 to *average*, 40 to *poor*, and 20 to *worst*. In this way, a distributed assessment can be transformed to a quality score. Finally, we ranked medical quality of the three studied hospitals on the basis of the computed combined quality scores.

3. Results

The main results generated from this study include distributed assessments and numerical quality scores of the three hospitals between 2006 and 2010 on the individual indicators, expert judgments, patient feedback, and combined medical quality. In the following presentation, we will skip individual indicators, expert judgments, patient feedback, and only show the overall or combined medical quality of each hospital.

The final distributed assessment results of the studied hospitals after data aggregation are shown in Table 6.

It is clear from Table 6 that each hospital did very well in some areas over the studied period but also had poor performances in other areas. Such a spread of performances provides an informative basis for the better communication of the assessment results and the identification of areas for future improvement with confidence. Based on the distributed medical quality assessment results, we also computed the combined quality score for each hospital. The medical quality trend and ranking of the studied hospitals from 2006 to 2010 are shown in Fig. 2.

From Fig. 2, we can see that the yearly quality ranking changed during the studied period, and the medical quality of the three hospitals had significant differences in 2006, but the differences were very small in 2010 after four years of development. We showed the results to the managers of the three hospitals, and they found that the results were very close to reality.

4. Discussion and conclusions

The objective of this study is to use the ER approach to aggregate multiple objective quality indicators, subjective expert judgments and patient feedback in a hierarchical structure to produce the combined overall assessments of quality of care of hospitals, so as to rank hospitals based on quality and provide an informative yet rigorous basis for quality improvement. The quality assessment framework employed in this study is based on Donabedian's

Table 6
Combined medical quality assessments of Hospitals A, B, and C (2006–2010).

Hospitals	Year	Distributed assessment after data aggregation
Hospital A	2006	{(excellent, 13.97%), (good, 31.64%), (average, 38.45%), (poor, 12.48%), (worst, 3.47%)}
	2007	{(excellent, 10.10%), (good, 35.48%), (average, 36.87%), (poor, 15.97%), (worst, 1.58%)}
	2008	{(excellent, 10.54%), (good, 40.98%), (average, 32.77%), (poor, 13.43%), (worst, 2.28%)}
	2009	{(excellent, 10.09%), (good, 49.28%), (average, 24.98%), (poor, 12.93%), (worst, 2.73%)}
	2010	{(excellent, 16.46%), (good, 41.84%), (average, 29.85%), (poor, 11.30%), (worst, 0.55%)}
Hospital B	2006	{(excellent, 9.04%), (good, 42.04%), (average, 24.38%), (poor, 20.38%), (worst, 4.16%)}
	2007	{(excellent, 12.28%), (good, 45.30%), (average, 26.33%), (poor, 8.87%), (worst, 7.23%)}
	2008	{(excellent, 10.03%), (good, 48.62%), (average, 31.15%), (poor, 8.61%), (worst, 1.60%)}
	2009	{(excellent, 13.12%), (good, 50.99%), (average, 25.80%), (poor, 9.29%), (worst, 0.79%)}
	2010	{(excellent, 10.34%), (good, 52.76%), (average, 27.06%), (poor, 7.45%), (worst, 2.38%)}
Hospital C	2006	{(excellent, 24.56%), (good, 37.19%), (average, 29.16%), (poor, 5.60%), (worst, 3.49%)}
	2007	{(excellent, 21.69%), (good, 39.48%), (average, 26.82%), (poor, 8.86%), (worst, 3.16%)}
	2008	{(excellent, 16.67%), (good, 42.87%), (average, 25.10%), (poor, 10.74%), (worst, 4.62%)}
	2009	{(excellent, 16.28%), (good, 41.08%), (average, 29.27%), (poor, 11.68%), (worst, 1.68%)}
	2010	{(excellent, 14.85%), (good, 48.84%), (average, 26.74%), (poor, 8.38%), (worst, 1.19%)}

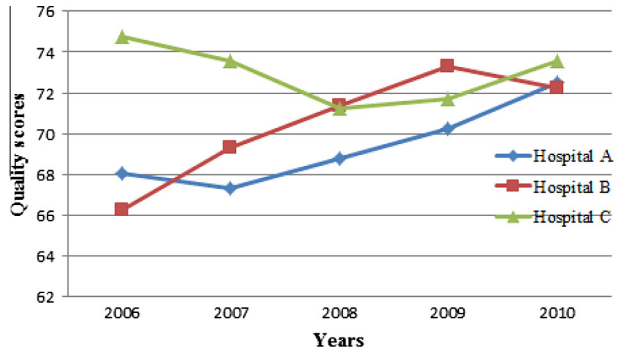

Fig. 2. Medical quality trend and ranking of Hospitals A, B, and C from 2006 to 2010.

quality model. In the ER-based quality evaluation approach, individual objective indicators, expert judgments and patient feedback on individual criteria were considered as pieces of evidence contributing to the overall quality of care, and these pieces of evidence were modeled using the same set of assessment grades within a belief structure. The recursive ER algorithm was used to combine multiple pieces of evidence one by one and layer by layer. Using the ER approach to generate combined quality assessment can help to provide a distributed assessment of quality on each alternative together with its contributing attributes or criteria, thereby providing a panoramic picture about quality at any level of concern to support informative decision making. In addition, the ER-based quality assessment results can help to provide an annual quality ranking together with a quality trend of each hospital during the studied period.

From a medical quality assessment perspective, this is the first time to systemically assess the quality of care by combining objective indicators, expert judgments and patient feedback, and the ER approach is employed for generating combined medical quality assessment for the first time as well. From an analytic methodology perspective, a pragmatic and useful method is proposed in this study for transforming numerical indicator values equivalently to qualitative assessment grades with a belief structure, and this helps to facilitate the combination of quantitative data and qualitative evaluations in the ER framework. The ER-based quality assessment method presented in this study can also give each hospital a fixed quality score for ranking, and the problem of rank reversal does not happen due to the employment of a common assessment scale at all criteria. From an expert system perspective, although user-friendly interfaces have not yet been developed, this study contributes to the medical quality research area with an intelligent system, where the inputs of the system can be medical dataset, surveyed dataset of expert judgments and patient feedback and the outputs of the system can be distributed quality assessments on different assessment grades together with the quality scores of hospitals. The core component of the system is the ER-based aggregation mechanism, and the criteria and the domain knowledge of the hierarchical quality assessment structure are embedded into the program.

Normally a distribution can show the diversity of the medical quality of a hospital. For hospital managers, they can identify poor areas to improve upon or good areas to maintain based on the distributed assessments. For patients, they can select a hospital that is excellent in the areas of their concern. Quality ranking among hospitals and a hospital's own quality trend can also play important roles in today's medical services. For governments all around the world, quality ranking among hospitals can be used as a criterion for medical resource allocation. For hospital managers, as the fluctuation or improvement of medical quality is closely related to hospital management policies, from the trend chart managers can examine which measures they took, and in which year, were effective in improving medical quality. For healthcare consumers-patients, quality ranking in different areas among peer hospitals can provide them with a direct reference for choosing the right hospitals to get high-quality medical services in the areas of their concern.

Inevitably, this study has limitations. Firstly, only a limited number of indicators can be derived from IMRSs. Secondly, it is difficult for experts or patients to provide clear retrospective judgments about the annual quality of the studied hospitals from 2006 to 2010. However, IMRSs are the only objective hard data that we can obtain and expert judgments and patient feedback are subjective in nature. Nevertheless, the medical quality assessment results generated by this study are probably the most objective and comprehensive on the basis of the available data for the studied hospitals. The credibility of such assessment results could be made even higher if more objective data, more timely and representative expert judgments and wider patient feedback become available.

In our future research, an intelligent medical quality assessment system will be developed, which could automatically extract quality indicators from available medical data sources such as IMRSs systems, acquire timely expert judgments and patient views via user-friendly interfaces and probably the Internet, and regularly produce combined quality assessment results. In addition, objective indicators and subjective evaluation criteria will need to be more flexible and scalable and can be defined by system users.

Acknowledgments

This study was supported by a humanities and social sciences project funded by the Ministry of Education of China under Grant No.: 13YJC630066. This study was also supported by a grant from the National Natural Science Foundation of China (NSFC) under Grant No.: 81301296.

References

Ashton, C., Kuykendall, D. H., Johnson, M. L., Wun, C. C., Wray, N. P., Carr, M. J., et al. (1994). A method of developing and weighting explicit process of care criteria for quality assessment. *Medical Care, 32*(8), 755–770.

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*(3), 2341–2354.

Büyüközkan, G., Çifçi, G., & Güleryüz, S. (2011). Strategic analysis of healthcare service quality using fuzzy AHP methodology. *Expert Systems with Applications, 38*(8), 9407–9424.

Baker, D. W. M. D. M. P. H., Einstadter, D. M. D. M. P. H., Thomas, C. L. B. A., Husak, S. S. B. S., Gordon, N. H. P., & Cebul, R. D. M. D. (2002). Mortality trends during a program that publicly reported hospital performance. *Medical Care, 40*(10), 879–890.

Bellows, N. M., & Halpin, H. A. (2008). Impact of medicaid reimbursement on mental health quality indicators. *Health Services Research, 43*(2), 582–597.

Benin, A. L., Vitkauskas, G., Thornquist, E., Shapiro, E. D., Concato, J., Aslan, M., et al. (2005). Validity of using an electronic medical record for assessing quality of care in an outpatient setting. *Medical Care, 43*(7), 691–698.

Bottle, A., & Aylin, P. (2008). Intelligent information: a national system for monitoring clinical performance. *Health Services Research, 43*(1), 10–31.

Campbell, S. M., Roland, M. O., & Buetow, S. A. (2000). Defining quality of care. *Social Science & Medicine, 51*(11), 1611–1625.

Carlucci, D., Renna, P., & Schiuma, G. (2013). Evaluating service quality dimensions as antecedents to outpatient satisfaction using back propagation neural network. *Health Care Management Science, 16*(1), 37–44.

Cebul, R. D., Love, T. E., Jain, A. K., & Hebert, C. J. (2011). Electronic health records and quality of diabetes care. *New England Journal of Medicine, 365*(9), 825–833.

Chae, Y. M., Kim, H. S., Tark, K. C., Park, H. J., & Ho, S. H. (2003). Analysis of healthcare quality indicator using data mining and decision support system. *Expert Systems with Applications, 24*(2), 167–172.

Chang, W.-J., & Chang, Y.-H. (2013). Patient satisfaction analysis: Identifying key drivers and enhancing service quality of dental care. *Journal of Dental Sciences, 8*(3), 239–247.

Contencin, P., Falcoff, H., & Doumenc, M. (2006). Review of performance assessment and improvement in ambulatory medical care. *Health Policy, 77*(1), 64–75.

Dijs-Elsinga, J., Otten, W., Versluijs, M. M., Smeets, H. J., Kievit, J., Vree, R., et al. (2010). Choosing a hospital for surgery: The importance of information on quality of care. *Medical Decision Making, 30*(5), 544–555.
Donabedian, A. (1966). Evaluating the quality of medical care. *The Milbank Memorial Fund Quarterly, 44*(3), 166–203.
Donabedian, A. (1968). Promoting quality through evaluating the process of patient care. *Medical Care, 6*(3), 181–202.
Farley, H., Enguidanos, E. R., Coletti, C. M., Honigman, L., Mazzeo, A., Pinson, T. B., et al. (2014). Patient satisfaction surveys and quality of care: an information paper. *Annals of Emergency Medicine, 64*(4), 351–357.
Feinstein, A. R. (2002). Is "quality of care" being mislabeled or mismeasured? *The American Journal of Medicine, 112*(6), 472–478.
Glance, L. G., Dick, A. W., Mukamel, D. B., Li, Y., & Osler, T. M. (2010). How well do hospital mortality rates reported in the New York state CABG report card predict subsequent hospital performance? *Medical Care, 48*(5), 466–471.
Glazer, J., McGuire, T. G., Cao, Z., & Zaslavsky, A. (2008). Using global ratings of health plans to improve the quality of health care. *Journal of Health Economics, 27*(5), 1182–1195.
Goodson, J., & Jang, W. (2008). Assessing nursing home care quality through Bayesian networks. *Health Care Management Science, 11*(4), 382–392.
Halfon, P. M. D. M. P. H., Eggli, Y. M. D. P., Pretre-Rohrbach, I. M. D., Meylan, D. M., Marazzi, A. P., & Burnand, B. M. D. M. P. H. (2006). Validation of the potentially avoidable hospital readmission rate as a routine indicator of the quality of hospital care. *Medical Care, 44*(11), 972–981.
Heineken, P. A., Charles, G., Stimson, D. H., Wenell, C., & Stimson, R. H. (1985). The use of negative indexes of health to evaluate quality of care in a primary-care group practice. *Medical Care, 23*(3), 198–208.
Hofer, T. P., & Hayward, R. A. (1996). Identifying poor-quality hospitals: can hospital mortality rates detect quality problems for medical diagnoses? *Medical Care, 34*(8), 737–753.
http://www.wechat.com/en/. WeChat for iPhone [Online]. [Accessed 15, February 2015].
Jiang, J., Li, X., Zhou, Z.-J., Xu, D.-L., & Chen, Y.-W. (2011). Weapon system capability assessment under uncertainty based on the evidential reasoning approach. *Expert Systems with Applications, 38*(11), 13773–13784.
Kerr, E. A., Hofer, T. P., Hayward, R. A., Adams, J. L., Hogan, M. M., McGlynn, E. A., et al. (2007). Quality by any other name?: A comparison of three profiling systems for assessing health care quality. *Health Services Research, 42*(5), 2070–2087.
Kipnis, P. P., Escobar, G. J. M. D., & Draper, D. P. (2010). Effect of choice of estimation method on inter-hospital mortality rate comparisons. *Medical Care, 48*(5), 458–465.
Kong, G. L., Ma, X. M., Zhao, L. P., & Zhang, J., 2013. Medical quality assessment using the evidential reasoning approach, INFORMS Healthcare Conference. Chicago, US.
Kong, G. L., Xu, D. L., Body, R., Yang, J. B., Mackway-Jones, K. R. H., & Carley, S. (2012). A belief rule-based decision support system for clincial risk assessment of cardiac chest pain. *European Journal of Operational Research, 219*(3), 564–573.
MacLean, C. H., Louie, R., Shekelle, P. G., Roth, C. P., Saliba, D., Higashi, T., et al. (2006). Comparison of administrative data and medical records to measure the quality of medical care provided to vulnerable older patients. *Medical Care, 44*(2), 141–148.
Marang-van de Mheen, P. J., Dijs-Elsinga, J., Otten, W., Versluijs, M., Smeets, H. J., Vree, R., et al. (2011). The relative importance of quality of care information when choosing a hospital for surgical treatment: A hospital choice experiment. *Medical Decision Making, 31*(6), 816–827.
Mcqueen, L., Mittman, B. S., & Demakis, J. G. (2004). Overview of the veterans health administration (VHA) quality enhancement research initiative (QUERI). *Journal of American Medical Informatics Association, 11*(5), 339–343.
Mull, H. J., Borzecki, A. M., Loveland, S., Hickson, K., Chen, Q., MacDonald, S., et al. (2014). Detecting adverse events in surgery: Comparing events detected by the veterans health administration surgical quality improvement program and the patient safety indicators. *The American Journal of Surgery, 207*(4), 584–595.
Normand, S.-L. T., Wolf, R. E., & McNeil, B. J. (2008). Discriminating quality of hospital care in the United States. *Medical Decision Making, 28*(3), 308–322.
Pearson, M. L., Lee, J. L., Chang, B. L., Elliott, M., Kahn, K. L., & Rubenstein, L. V. (2000). Structured implicit review: A new method for monitoring nursing care quality. *Medical Care, 38*(11), 1074–1091.
Ploeg, A. J., Flu, H. C., Lardenoye, J. H. P., Hamming, J. F., & Breslau, P. J. (2010). Assessing the quality of surgical care in vascular surgery; moving from outcome towards structural and process measures. *European Journal of Vascular and Endovascular Surgery, 40*, 696–707.
Pouwer, F., & Snoek, F. J. (2002). Patients' evaluation of the quality of diabetes care (PEQD): Development and validation of a new instrument. *Quality and Safety in Health Care, 11*(2), 131–136.
Rees, M., & Dineschandra, J. (2005). Monitoring clinical performance. The role of software architecture. *Health Care Management Science, 8*(3), 197–203.
Reeves, D., Campbell, S. M., Adams, J., Shekelle, P. G., Kontopantelis, E., & Roland, M. O. (2007). Combining multiple indicators of clinical quality: an evaluation of different analytic approaches. *Medical Care, 45*(6), 489–496.
Ritchey, J., Gay, E. G., Spencer, B. A., Miller, D. C., Wallner, L. P., Stewart, A. K., et al. (2012). Assessment of the quality of medical care among patients with early stage prostate cancer undergoing expectant management in the United States. *The Journal of Urology, 188*, 769–774.
Rosenthal, G. E., Shah, A., Way, L. E., & Harper, D. L. (1998). Variations in standardized hospital mortality rates for six common medical diagnoses: implications for profiling hospital quality. *Medical Care, 36*(7), 955–964.
Thomas, J. W., & Hofer, T. P. (1999). Accuracy of risk-adjusted mortality rate as a measure of hospital quality of care. *Medical Care, 37*(1), 83–92.
Untachai, S. (2013). Modeling service quality in hospital as a second order factor, Thailand. *Procedia - Social and Behavioral Sciences, 88*, 118–133.
Vasudevan, A., Arachchi, A., & van Langenberg, D. R. (2013). Assessing patient satisfaction in inflammatory bowel disease using the QUOTE-IBD survey: A small step for clinicians, a potentially large step for improving quality of care. *Journal of Crohn's and Colitis, 7*(9), e367–e374.
Wang, Y.-M., & Elhag, T. M. S. (2008). Evidential reasoning approach for bridge condition assessment. *Expert Systems with Applications, 34*(1), 689–699.
Wang, Y. M., Yang, J. B., & Xu, D. L. (2006). Environmental impact assessment using the evidential reasoning approach. *European Journal of Operational Research, 174*(3), 1885–1913.
Weissman, J. S., Ayanian, J. Z., Chasan-Taber, S., Sherwood, M. J., Roth, C., & Epstein, A. M. (1999). Hospital readmissions and quality of care. *Medical Care, 37*(5), 490–501.
Wray, N. P., Peterson, N. J., Souchek, J., Ashton, C. M., & Hollingsworth, J. C. (1997). Application of an analytic model to early readmission rates within the department of veterans affairs. *Medical Care, 35*(8), 768–781.
Xu, D. L. (2009). Assessment of nuclear waste repository options using the ER approach. *International Journal of Information Technology & Decision Making, 8*(3), 581–607.
Xu, D. L., McCarthy, G., & Yang, J. B. (2006). Intelligent decision system and its application in business innovation self assessment. *Decision Support Systems, 42*(2), 664–673.
Yang, J. B. (2001). Rule and utility based evidential reasoning approach for multiple attribute decision analysis under uncertainty. *European Journal of Operational Research, 131*(1), 31–61.
Yang, J. B., & Singh, M. G. (1994). An evidential reasoning approach for multiple-attribute decision making with uncertainty. *IEEE Transactions on Systems, Man and Cybernetics, 24*(1), 1–18.
Yang, J. B., & Xu, D. L. (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*(3), 289–304.