Abstract
Introduction: Selecting the best set of input and output indicators and allocation of correct weights to them is a sensitive step in any efficiency evaluation study. Therefore, the present study aims to determine and rank the efficiency indicators of hospitals.

Methods: This mixed-method study was carried out in three steps: comprehensive literature review, application of the Delphi method to determine the best indicators for efficiency evaluation of the hospitals, and utilization of a fuzzy analytic hierarchical process (FAHP) for weighting of final indicators and ranking them.

Results: 8 input and 9 output indicators were selected for efficiency evaluation of the hospitals which were weighted by FAHP. Among the input indicators, the number of physicians and active beds and among the output indicators, length of stay and number of surgeries were identified to be the most important indicators.

Conclusion: According to the proposed indicators and their accurate weights, efficiency evaluation of hospitals can be done more accurately, reliably, and comprehensively.

Keywords: Efficiency, Indicator, Delphi, Fuzzy Analytic Hierarchical Process, Hospital.

Introduction
Performance evaluation of different units has been a primary concern of managers for a long time. The scientific approach to this subject has started and expanded significantly since the World War II (1, 2). In this regard, it is important to evaluate the performance of hospitals that provide healthcare services and training of the professionals and are considered as the costliest part of the health sector (1-3). In the performance evaluation of hospitals, the main incentive for the utilization of functional and scientific methods is the optimal consumption of physical, technological, and human resources. (4). Efficiency evaluation is considered to be the first step of performance evaluation that provides a logical framework for the distribution of human and financial resources (5).

Efficiency evaluation can be used for discovering the best ways to improve productivity and efficiency, do the necessary reforms, and allocate the resources efficiently (6). Efficiency refers to the ability of hospitals in using their resources for the best productions and services in comparison with their similar rivals. It also refers to the number of hospital productions or outputs that can be increased without any changes in their inputs (7, 8).

Moreover, efficiency has a great significance as a part of productivity. Efficiency estimation is an indicator for performance evaluation of similar and homogenous units (9). By efficiency evaluation, efficient hospitals can preserve or improve their efficiency by identifying their strengths and weaknesses and by using the experiences of superior units. Also, inefficient hospitals can become closer to efficiency borders with benchmarking of the reference units and better management of their inputs (10, 11).

Determining and selecting the best set of input and output indicators is one of the most sensitive steps in
efficiency evaluation (12, 13); accurate determination of input and output indicators leads to more real efficiency evaluation (14, 15). Moreover, one of the weaknesses of the efficiency evaluation techniques is the uncontrolled final weights which are calculated for inputs and outputs. In other words, since the final weights are calculated without judgment of the decision-makers, they may show an almost efficient unit, while this efficiency is not intrinsic and only results from its weight (16).

In this regard, most of the previous studies have not used a specific method to select and determine the most appropriate indicators for hospital efficiency evaluation and almost all have used similar indicators (1, 8-14, 17, 18-30).

Against the quantitative methods, qualitative methods have recently been taken into consideration; their general characteristic is that the weights resulting from these methods are dependent on experience and judgment of the experts. It has the advantage of defining weight priorities based on the real importance of each indicator. Therefore, there will be no contrast between the weight of indicators. Practically, using each of the qualitative or quantitative methods separately cannot reveal the natural differences between the weights of the indicators. In this way, the efficiency value will be questioned. Therefore, the advantages of both qualitative and quantitative methods should be combined to increase the accuracy and reliability of future evaluations.

Methods
This mixed-method study was done in the following three phases:

Phase 1: Conducting a comprehensive literature review for identification of the efficiency indicators
Initially, a comprehensive literature review of previous studies was done regarding the efficiency evaluation of hospitals and then a complete list of input and output indicators that were suitable for efficiency evaluation of hospitals was prepared for use in the next phase of the study.

For the literature review, the keywords such as “efficiency+hospital”, “DEA+hospital”, and “data envelopment analysis+hospital” were used to search in article titles in different databases (PubMed, Science Direct, ProQuest, Cochrane, and Scopus for articles written in the English language and Google, MAGIRAN, and SID for articles in the Persian language) from 2000 to 2016. This search was done in October 2016.

Phase 2: Conducting a Delphi technique for acquiring the consensus
In this step, the Delphi technique was used for the selection of the best indicators extracted from the papers and of experts’ opinions for efficiency evaluation of the hospitals. The members of Delphi included faculty members of health care management, Hospital managers of Shiraz University of Medical Sciences, and Ph.D. candidates of health care management in Shiraz University of Medical Sciences (total numbers=15) which were selected through purposeful sampling. The inclusion criteria for Delphi specialists in this study were their willingness to collaborate in the research and having related knowledge or previous research in this field. Also, the exclusion criteria were the absence of the two above cases. There was a Likert scale ranging from 1 to 9 (10 to 90 percent) for calculation of the consensus at the appropriateness of the input and output indicators for efficiency evaluation of the hospitals. In this study, Delphi was run in three rounds. In the first round, experts were also asked to complete the list of indicators, if necessary, by adding their own to the existing list. In each round, those indicators which gained less than 40% agreement were removed from the indicators list. Those indicators which gained a consensus between 40 to 60% agreement entered the next round, and those with more than 70% (strong) agreement were selected as the best indicators for efficiency evaluation of the hospitals.

Phase 3: Conducting FAHP for weighting the selected indicators
In this step, the selected indicators of the first phase were weighted, using FAHP. The expert members that were considered to be panel members of the Delphi technique in the previous step participated in FAHP paired comparisons of this step, too.

The first step in the prioritization of decision-making elements is paired comparisons, i.e. a comparison of elements in pairs according to specific criteria (21). Thus, in this step, the final indicators
which resulted from the Delphi technique (8 input and 9 output indicators) were used as a checklist for paired comparisons. In this checklist, indicators were placed two by two in front of each other and a descriptive spectrum was considered between them. This spectrum included: “exactly the same”, “similar (or very weak)”, “weak”, “strong”, “very strong”, and “absolute”. Options of this spectrum indicated the importance of each indicator for efficiency evaluation of hospitals. After designing a paired comparison checklist, it was sent to the experts.

During entering data into the software (Excel), linguistic options of the FAHP checklists, which were filled by experts, as shown in Table 1, were converted to triangular fuzzy numbers, so that three triangular fuzzy numbers were allocated to each linguistic option. For each checklist completed by experts (n=15), an 8x24 matrix was formulated for input indicators and a 9x27 matrix for output indicators. If the first indicator was superior to the second one in the FAHP checklist, the amount of its superiority was determined by one of the experts based on the second row of Table 2 in the form of three fuzzy numbers; However, if the second indicator was superior to the first one, the amount of its superiority was indicated by the reverse of three fuzzy numbers based on the third row of Table 1.

After formulating all the matrices of FAHP checklists, two total matrices were created for input and output indicators. Each block of these matrices included the geometric means of all the blocks corresponding to that specific block in other matrixes. The new total matrix was calculated by adding all the columns and rows of the total matrix and multiplied in a reverse matrix to calculate the final decision-making matrix. Using the final decision-making matrix, in each stage, the score of each indicator was compared with other indicators.

Finally, according to paired comparisons, multiple scores were obtained for each indicator in which the lowest score is considered to be the weight of that indicator. In the final step, the obtained weights were normalized and the final input and output weights were calculated; the sum of all these weights was equal to number one.

**Results**

Figure 1 shows the flowchart of the study selection. As shown in this Figure, 126 Persian and 204 English papers were found in our search; after checking the papers’ titles and removing the repetitive ones, 111 English and 45 Persian papers entered the abstract and

| Table 1: Fuzzy triangular numbers for the corresponding linguistic options of FAHP checklist |
|-----------------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| The spectrum of the importance degree of linguistic options | Exactly the same | Similar | Weak | Strong | Very Strong | Absolute |
| Main fuzzy triangular numbers                  | 1 | 1 | 1 | 0.5 | 1 | 1 | 1.5 | 1.5 | 2 | 1.5 | 2 | 2.5 | 2 | 2.5 | 3 | 2.5 | 3 | 2.5 | 3 | 2.5 | 3 | 3.5 |
| Reverse of fuzzy triangular numbers             | 1 | 1 | 1 | 0.666 | 1 | 2 | 0.5 | 0.666 | 1 | 0.4 | 0.666 | 0.333 | 0.4 | 0.5 | 0.285 | 0.333 | 0.4 |

| Table 2: Final input and output indicators for efficiency evaluation of the hospitals selected by Delphi technique |
|---------------------------------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Input Indicators                                              | Output Indicators |
| Number of active beds                                         | Average outpatient and emergency visits each month |
| Number of full-time general physicians                       | Average number of hospitalized patients in each month |
| Number of specialists                                         | Average number of surgeries in each month |
| Number of nurses                                              | Average length of stay of each patient |
| Number of service providers except doctors and nurses         | Bed occupancy rate |
| Monthly average of all the executive costs of hospital       | Average bed turnover rate |
| Monthly average costs of workers and fees                     | Average death rate in each month |
| Hoteling average costs                                       | Average discharge rate with patient’s agreement |
| Monthly average costs                                         | Average income of hospital in each month |
In this literature review, 39 input and 52 output indicators were extracted. The most common input indicators among all the papers were the number of beds, number of physicians and number of nurses with frequencies of 44, 38, and 25, respectively. Also, the most common output indicators were the number of outpatients and emergency patients, number of inpatients, number of surgeries, bed occupancy rate, and length of stay with frequencies of 36, 28, 16, 14, and 12, respectively.

The extracted indicators were entered into the first Delphi round, in which those with less than 40% agreement were removed from the list (5 and 10 indicators, respectively); those with 40 to 60% agreement (17 and 7 indicators, respectively) were entered into the next round and indicators which had gained more than 70% agreement were selected to be included in the final list. Eventually, after the third round of the Delphi technique, 17 indicators, including 8 input and 9 output indicators, were selected for efficiency evaluation of the hospitals.
These indicators were confirmed by experts (Table 2).

Table 3 indicates the relative (un-normalized) and normalized weights of all the input and output indicators which were calculated via dividing each relative weight by their sum. Consequently, the sum of the final weights of input indicators and output indicators was calculated to be equal to one.

As presented in Table 3, the most important input indicators are the number of physicians, active beds, and number of nurses, respectively. Also, the most important output indicators are the length of stay, number of surgeries, and occupancy rate, respectively.

Discussion
This study aimed to identify and rank the efficiency indicators for hospitals. A total number of 91 input and output indicators were identified from the literature review. The most common input indicators were the number of hospital beds, physicians and nurses, and the most common output indicators were the number of outpatient and emergency patients, hospitalized patients, surgeries, bed occupancy rate, and average length of stay.

Regarding the frequency of these indicators in different studies, a systematic review study by O’Neill et al. (2008), which was conducted on 79 efficiency indicators of hospitals during 1984-2004, indicated that the most common input indicators used in different studies are the number of active beds (in 55 studies) and number of clinical personnel (in 54 studies) including physicians, nurses, and service providers. In some of these studies, input indicators of physicians and nurses were divided into subcategories of specialist, general physician, etc. Moreover, O’Neill’s study concluded that the most common output indicators are outpatient visits (in 52 studies), inpatient cases (in 18 studies), and number of surgeries (in 22 studies) (6).

In this respect, Rahimi (2014) conducted a systematic review of the indicators of hospital performance evaluations. He concluded that the bed occupancy rate and length of stay with frequencies of 13 and 10, respectively, were the most common indicators (22). These results are consistent with those of the present study. Safari (2010) analyzed various financial and non-financial indicators in his research and identified 27 financial and general indicators (23).

The final 8 input and 9 output indicators which resulted from two rounds of Delphi technique were confirmed by the experts. The input indicators included the number of active beds, number of full-time general physicians, number of specialist physicians, number of nurses, number of other personnel except for the doctors and nurses, monthly average of total operating costs of the hospital, monthly average of the staff costs (salaries), and average of hoteling costs. The output indicators included the average of outpatient and emergency visits in each month, monthly average number of inpatients, monthly average number of surgeries, average length of stay, bed occupancy rate, average of bed turnover, monthly mortality rate, average of discharges against medical advice, and average of monthly hospital income (2012). Safari (2010), Azar (2013), Parvizian (2012), and Kashani Pour (2008) also reported the same set of indicators for efficiency evaluation of the hospitals (23-26).

Also, in the study of Mateus (2015), hospital discharge was used as the output indicator and the number of staff, nurses, physicians, and hospital beds were considered as the input indicators for calculating the efficiency of hospitals in four European countries (Portugal, Spain, Slovenia, and the United Kingdom) which is in agreement with the present study (27).

Although Saleh Zadeh (2011), Rashidian (2010), Pour Reza (2009), and Ilbeygi (2012) made data envelopment analysis and suggested that some indicators, such as treatment quality, satisfaction of patients and personnel, hospital infection control indicators, and case-mixes can be taken into consideration for efficiency or performance evaluation of hospitals (28-31), such indicators were not approved by experts in this study to be used in efficiency evaluations of hospitals.

In the present study, FAHP was used for weighting and prioritizing the input and output indicators. According to the FAHP experts’ opinion, the most important input indicators were the number of physicians, active beds, and nurses, and the most important output indicators were the length of stay, number of surgeries, and bed occupancy rate. In the study of Saleh Zadeh (2011), two most important input indicators were the number of nurses and active beds (28). Similarly, in studies of Alem Tabriz (2010), Kashani Pour (2008), Ghasemi (2011), Azadeh (2009), Che (2010), and Tseng (2009), AHP method was used for identification of input and output indicators for efficiency evaluation of the desired units (9, 26, 32-35).

Conclusion
This study proposes a set of appropriate efficiency indicators for hospital evaluations. Using these indicators allows the efficiency evaluation to be done more accurately and comprehensively which, in turn, leads to better managerial decisions regarding resource allocation.
Authors’ contribution
NH designed the study and its overall methodology. NH also edited and finalized the article. TSH searched the databases and with the help of MKRZ retrieved the sources and scanned and screened all the articles. TSH also prepared the draft of the article. TSH and MKRZ contributed to data analysis and synthesis. Also, the study was under consultation and supervision by ZK and as advisor. All the authors have read and approved the final manuscript. Also, the study was approved by the ethical committee of Shiraz University of Medical Sciences (Ethical code of project).

Acknowledgments
The authors wish to express their appreciation to all the hospitals’ authorities who participated in this research. This study was a part of an approved research project at Shiraz University of Medical Sciences with the code number of 7120-93. Ethical code of project: IR.SUMS.REC.1393.S7120.

Conflict of Interest: None declared.

References
18. Bahrami MA, Rafiei S, Abedi M, Askari R. Data...


