# Educational ranking of medical universities in Iran (ERMU)

AliAkbar Haghdoost[1], Nader Momtazmanesh[2], Farangis Shoghi Shafagh Aria[3], Hadi Ranjbar*[4]

## Abstract

**Background:** Several international ranking systems are available for examining the performance of universities internationally. However, it seems that in these rankings the educational functions are relatively discounted. In this study, we presented the theoretical framework, methodology, and results of 2 rounds of ERMU (educational ranking of medical universities) in Iran.

**Methods:** The indicators were designed through literature reviews, expert panels, and consensus among professionals using Delphi method. Six domains, including quantitative expansion, educational management scheme, educational management, qualitative development, and alignment with the National Master Plan for Science and Education were assessed. Nineteen educational processes and outcome-based indicators were defined at the end. Some of them, such as the number of educational disciplines, were normalized based on the number of academic staff. Using exploratory factorial analysis, the interval validity of the final instrument was assessed. Also, the internal consistency of the tool was evaluated by calculating Cronbach's alpha. Data were collected in 2 rounds in 2010 and 2014. The required data were collected and cross-validated from different sources in medical universities and in the Ministry of Health.

**Results:** The analysis yielded 2 factors, explaining a total of 61.6% of the variance for the entire set of variables. The calculated Cronbach's alpha was 0.94. The correlation coefficient for all items was 0.88. The comparison of the results of the 2 rounds showed a significant improvement in the scores of universities in all domains.

**Conclusion:** The results showed a relatively fitted model with acceptable reliability. In addition, it seems that the ranking of medical universities has improved their performance in the study indictors. It seems that in the long-term, monitoring these indicators and the visibility of their scores may improve the educational performance of universities.

**Keywords:** National, University, Ranking, Higher education, Academic performance

## Introduction

Benchmarking and performance analysis of universities has gained increasing interest over the past few decades (1) and has become a national and international competition (2). The first rankings were performed at the end of

*Corresponding author: Dr Hadi Ranjbar, ranjbar.h@iums.ac.ir*

1. Modeling in Health Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran
2. Shahid Beheshti University of Medical Sciences, Tehran, Iran
3. Iran Ministry of Health and Medical Education, Tehran, Iran
4. Mental Health Research Center, Tehran Institute of Psychiatry- School of Behavioral Sciences and Mental Health, Iran University of Medical Sciences, Tehran, Iran

↑*What is "already known" in this topic:*
There are several international ranking systems that evaluate the performance of universities. Most ranking systems are based on research performance. Evaluation and ranking of education systems is an effective method to maximize the efficiency of their performance and to encourage them to achieve national objectives.

→*What this article adds:*
In this study, it was shown that a simple tool can be designed to assess the educational performance of medical universities and explore the impact of such an assessment on their performance over time. Also, it was discussed that this type of monitoring encourages the universities to work more efficiently to achieve the national higher education goals.

the 19th century, but modern ranking systems were performed at the end of the 20th century (3).

In 1983, USNWR (U.S. News and World Report) published the results of the first ranking of academic institutes in the United States. This was a national ranking that only assessed the universities in the United States. The beginning of the 21st century saw the introduction of international rankings. The most well-known international ratings include the Academic Ranking of World Universities by Shanghai Ranking, the QS World University Rankings, Times Higher Education, the Webometrics Ranking of World Universities, and NTU (National Taiwan University) Ranking. Each of these rankings has unique features (4, 5).

Among the new age performance analysis systems, the Academic Ranking of World Universities was the first world university ranking and was first published in 2003 (6). It ranks the world's top 1000 colleges and universities based on objective but research-based indicators. In 2010, Times Higher Education magazine announced a new method in collaboration with Thomson Reuters. Their 5 main indicators are industry income (innovation), international diversity, teaching (the learning environment), research (volume, income, reputation), and citations (research influence). The individual criteria for teaching are reputational survey (teaching), Ph.Ds.' awarded per academic year, income per academic year, and undergraduate degrees awarded per academic year (7, 8).

The Webometrics Ranking of World Universities is a ranking system for the world's universities based on the volume of their web content and the visibility and impact of these web publications. This ranking system was developed by the Cybermetrics Lab in Spain and ranking is done every 6 months and provides web indicators for more than 12000 universities worldwide. NTU Ranking provides overall ranking, rankings by 6 fields, and rankings by 14 selected subjects. The main focus of all the presented models is on publications by academic researchers.

Despite the development of numerous models of rankings and competition among them, each of these models has its own limitations and critics (9, 10). A general criticism of all these models is their lack of sufficient attention to the primary responsibility of universities, which is educating students.

Assessment of the quality of education is so complex that it has either been abandoned or alternative criteria has been used to measure it. Indicators, such as the number of faculty, the composition of graduates, and even the number of citations, are used as proxies for assessing the performance of education systems. Nevertheless, in all international rankings models, there is a greater emphasis on research. Indicators, such as winning the Nobel Prize, are inaccessible for most universities. Also, only a small number of universities manage to gain a place in international rankings, and the positions of many institutions at the national level are not specified. For these reasons, several countries have started to design their own ranking systems in recent years (1, 11-14).

Universities of Medical Sciences (UMS) in Iran are unique from a managerial point of view, as the health care services are being integrated into the educational and research systems within a defined catchment area. In other words, UMS are organizations that cover both scientific and care services. Due to this unique integrated system, the criteria and indicators for the monitoring and evaluation of their performance have to be formulated accordingly.

Based on the above explanation, we created a national framework to rank UMS in Iran to check their educational performance. Such a ranking system not only ranks UMS but also motivates them to work more intensively towards achieving national goals systematically. In this paper, we present the steps that were taken to develop the ranking indicators and the results of 2 rounds of ranking.

## Methods

ERMU (Educational Ranking of Medical Universities) ranking includes 54 medical universities in Iran. Data were collected in 2 rounds in 2010 and 2014. The universities uploaded their information using an online data collection form. Their records were then cross-validated with the registered data in databanks in the Health Ministry, and any discrepancies were explored and rechecked with the universities accordingly.

### Formulating indicators

The process of formulating indicator was conducted between 2009 and 2010. This process was supervised and organized by a group of 15 experts in the field of medical education. The inclusion criteria for the selection of experts was management experience in medical education and a deep understanding of the educational system of Iran. The Delphi method was used to collect the opinions of experts across the country and create a consensus.

Primary indicators were developed through a systematic review and an extensive desk review. Only those indicators that met the following criteria were then included: (1) simple and reliable, (2) representative of the main educational activities, (3) relevant to educational processes within the UMS, (4) stable over time, and (5) conceptually have maximum consistency with the other indicators.

A comprehensive consensus was reached after several rounds of expert panels over 6 months of negotiation. In the next step, the operational definitions of indicators were finalized using 2 rounds of virtual discussion using the Delphi method.

The primary draft of the data collection form was distributed to all vice-chancellors for education across UMS in Ira, and their feedback was systematically collected in 2 phases: (1) face to face, (2) via email and official communications. The final version of indicators and data collection forms were then approved in a national assembly of vice-chancellors of education. To confirm the feasibility of the indicators, the research team held a meeting with faculty and education managers from one of the medical universities and their comments about feasibility were applied. Finally, 19 indicators and their weights in 5 domains were approved. The domains, indicators, related weights, and operational definitions of each indicator are

*Table 1.* The definition and weight of domains and indicators

| Domain | W | Index | Definition | W |
|---|---|---|---|---|
| Quantitative expansion | 20 | Newly established disciplines and degrees | The number of academic disciplines and degrees that were established in the past three years | 60 |
| | | Existing discipline and degrees | The number of existing academic disciplines and degrees prior to expansion | 40 |
| The educational management scheme | 25 | Long and short term programing and performance reporting | Having approved educational strategic and operational plans Publishing a performance report annually | 10 |
| | | The role and performance of the university council | The number of University Council meetings Legislation related to education | 15 |
| | | The main formal coordination of educational activities | Educational council legislation performance | 15 |
| | | The starring role of the educational commission | Involving the schools in education legislation | 15 |
| | | Faculty hiring committee | The accuracy of hiring processes of faculty members | 10 |
| | | Educational resource allocation | Collaborative process of resource allocation | 15 |
| | | Management stability and expertise | Stability in use of expert managers in administrative positions in the field of education | 20 |
| Educational management | 20 | Education office website | The quality of content of the university website | 30 |
| | | Standard deviation of special courses scores | The standard deviation of scores within special courses | 35 |
| | | Postgraduate education entrance | The percentage of graduated students who are accepted into postgraduate exams | 35 |
| Qualitative development | 20 | Education development center performance | | 50 |
| | | Continuous education | The number of continuous education courses in the university | 20 |
| | | Mutual relationship with ministerial directors | The ministerial directors point of view about performance at the university level | 15 |
| | | National medical education congress | The presence of the university in the national medical congress Educational processes Ranking in student Olympiad | 10 |
| | | Medical education special activities | Innovative educational process in the field of medical education | 5 |
| Alignment with the national master plan for science and education | 15 | Implementation of the program | Number of international agreements Number of international publications Number of international professors Number of International Students | 50 |
| | | Enhancing professionalism | Having a comprehensive medical ethics program Medical ethics committee performance Continuous education courses in professionalism Having an independent ethics department | 50 |

presented in Table 1.

### Instrument validity

Instrument validity was calculated using a quantitative method. The indicators were sent to vice chancellors for education of medical universities, who were asked to rate each item on a Likert scale from 1 to 5 for importance, measurability, precision, and improvability.

### Adjustment of scores based on the size of the UMS

As large universities might get higher scores in some indicators due to their stronger infrastructure, achievements were adjusted based on the number and the rank of their academic staff. This adjustment was only applied to the score of indicators 1, 2, and 14.

### Statistical analysis

The first author entered the data into SPSS 16 for analysis. Internal consistency was calculated for each domain and for all the items in the model. Exploratory factor analysis was used to examine the basic structure of the model. Also, principal component analysis was performed using the Quartimax rotation with Kaiser Normalization. The relationship between the scores from the 2 rounds was calculated using a Spearman's correlation. A

paired samples t test was performed to determine whether there was a statistically significant difference between the results of the 2 rounds (Table 1).

## Results

### Instrument validity part I: Content validity

The mean for each indicator is presented in Table 2. Enhancing professionalism, long- and short-term programming and performance reporting, and the main formal coordination of educational activities had the highest scores in terms of importance. In terms of measurability, the existing disciplines and degrees, newly established disciplines and degrees, and the main formal coordination of educational activities received the highest scores. Existing disciplines and degrees, newly established disciplines and degrees, and educational resource allocation scheme received the highest scores in terms of precision. The starring role of the educational commission, Education Development Center (EDC) performance, education office website, and existing disciplines and degrees had the highest scores in improvability.

### Instrument validity part II: Construct validity

The analysis yielded 2 factors explaining a total of 61.6% of the variance for the entire set of variables. Fac-

*Table 2.* The means of the importance, measurability, precision, and improvability of all indicators

| Domain | Index | Importance | Measurability | Precision | Improvability |
|---|---|---|---|---|---|
| Quantitative expansion | Newly established disciplines and degrees | 3.97 | 4.84 | 4.88 | 3.50 |
| | Existing discipline and degrees | 4.50 | 4.91 | 4.94 | 4.23 |
| The educational management scheme | Long and short term programming and performance reporting | 4.84 | 3.66 | 3.72 | 4.03 |
| | The role and performance of the university council | 4.48 | 4.35 | 4.06 | 3.90 |
| | The main formal coordination of educational activities | 4.76 | 4.57 | 4.33 | 4.30 |
| | The starring role of the educational commission | 4.58 | 4.29 | 4.16 | 4.35 |
| | Faculty hiring committee | 4.48 | 3.27 | 3.37 | 4.03 |
| | Educational resource allocation | 4.52 | 4.52 | 4.61 | 3.68 |
| | The experience and skills of the managerial team in the education sector | 4.26 | 4.38 | 4.47 | 3.69 |
| Educational endpoints | Education office website | 4.53 | 4.10 | 3.87 | 4.23 |
| | Internal consistency of students' scores in professional courses | 4.25 | 4.00 | 4.00 | 3.77 |
| | Postgraduate education entrance | 4.32 | 3.97 | 4.03 | 3.59 |
| Qualitative Advance | Education development center performance | 4.69 | 4.03 | 4.06 | 4.34 |
| | Continuous education courses for graduated personnel | 4.58 | 4.50 | 4.13 | 3.97 |
| | External evaluation of performance of education system in medical universities | 3.79 | 3.37 | 2.82 | 3.36 |
| | Nationally nominated activities of universities in annual medical education festivals | 4.26 | 4.22 | 3.88 | 3.94 |
| | Educational innovations | 4.25 | 3.97 | 3.84 | 4.19 |
| Alignment with the National Master Plan for Science and Education development | Internationalization | 4.71 | 3.50 | 3.44 | 3.81 |
| | Professionalism | 4.87 | 3.35 | 3.30 | 4.10 |

Each domain scored between 1 to 5 in columns

*Table 3.* Factor analysis of ranking indicators

| | Components | |
|---|---|---|
| | Process related indicators | Outcome indicators |
| Newly established disciplines and degrees | 0.911 | |
| Existing disciplines and degrees | 0.886 | |
| Internationalization | 0.884 | |
| Continuous education courses for graduated personnel | 0.867 | |
| External evaluation of performance of education system in medical universities | 0.850 | |
| Professionalism | 0.830 | |
| Long and short term programming and performance reporting | 0.782 | |
| Education office website | 0.769 | |
| Educational resource allocation | 0.729 | |
| Education development center performance | 0.701 | 0.423 |
| Nationally nominated activities of universities in annual medical education festivals | 0.696 | |
| Educational innovations | 0.670 | |
| Faculty hiring committee | 0.657 | |
| The experience and skills of the managerial team in the education sector | 0.635 | |
| The role and performance of the university council | 0.634 | |
| The starring role of the educational commission | 0.591 | |
| The main formal coordination of educational activities | 0.572 | |
| Postgraduate education entrance | | 0.765 |
| Internal consistency of students' scores in professional courses | | 0.711 |

tor 1 was labelled "process related indicators" and explained 52.11% of the variance. The second factor was labelled "outcome indicators" as it consisted of 2 indicators (standard deviation of scores of students in their courses, the success percentage of students to enter a higher level in postgraduate education) and explained 9.5% of the variance. The Kaiser-Meyer-Olkin measure of sampling adequacy showed a value of 0.845, and the Bartlett's test of sphericity showed a p-value less than 0.001. The factor analysis of indicators is presented in Table 3.

### Instrument reliability I: Internal consistency
The calculated alpha was 0.94 for the instrument, 0.95 for quantitative expansion, 0.83 for educational administration, 0.50 for educational management, 0.87 for qualitative development, and 0.89 for alignment of performance of universities compatible with the National Road Map for the Advancement of Science and Education.

### Instrument reliability II: Test retest
The results of the indicators and domains between the 2 rounds of data gathering were significantly correlated. The correlation coefficient for all of the items was 0.88. The correlation coefficient was 0.76 for quantitative expansion and 0.55 for educational administration. Educational management had a 0.60 correlation coefficient and qualitative development a correlation coefficient of 0.75. The correlation coefficient for alignment of performance of universi-
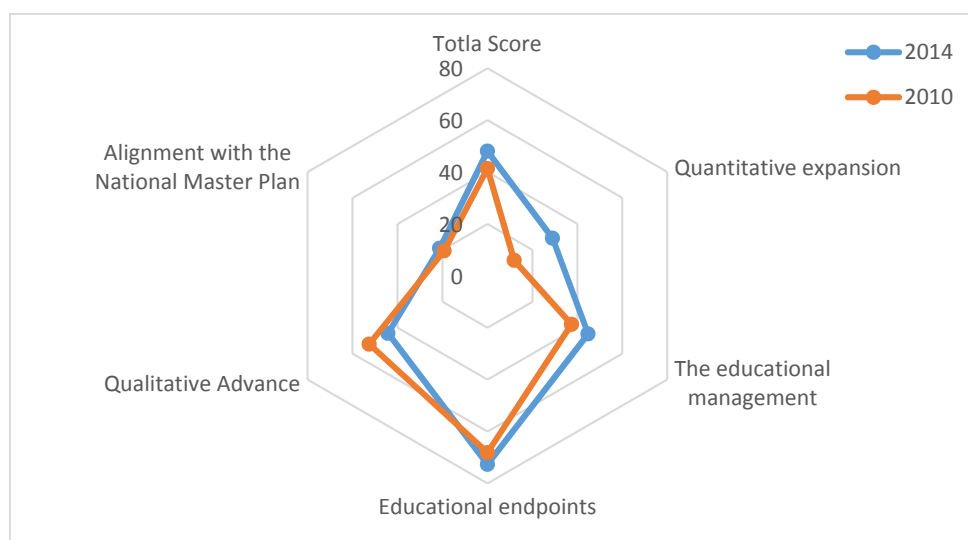
*Fig. 1.* Radar diagram of the five main components and the total scores from the 2010 round vs. the 2014 round

ties compatible with the National Road Map for the Advancement of Science and Education was 0.69.

### Comparison of the 2 rounds

The comparison of the results of the 2 rounds showed a significant increase in the scores of universities in all domains. The results in Figure 1 show that the difference between the 2 rounds was statistically significant in quantitative expansion (Z= -5.72, p< 0.001), educational administration (Z= -3.55, p< 0.001), educational management (Z= -2.72, p = 0.006), and total scores (Z= -5.44, p<0.001), with a higher score in the second round. The score for alignment with the National Road Map in the second round was higher, but the difference was not significant (Z= -1.49, p= 0.13). The score for qualitative development was lower in the second round, but the difference was not significant (Z= -1.87, p= 0.06).

### Discussion

The aim of this study was to report the process of formulating indicators and the results of 2 rounds of a national ranking system. Our results showed a high internal consistency and goodness of fit of the models. It also showed that there was a significant relationship between the results from the 2 rounds, with a significant increase in scores in the second round. Also, our results revealed a suitable internal consistency among ranking indicators.

There are several ranking models at international level with research-based indicators, while there are few national ranking models to assess the educational performance of universities. For these reasons, there have been numerous criticisms of the current ratings (4, 15, 16).

The international scope of the rating systems is one of the issues that has been the subject of many critiques. The number of institutions that can be surveyed in the ratings is so low that the vast majority of universities do not have the chance to be among them (11-13, 17). In these cases, the weight or importance of research indicators is also over presented in most of these methods. One of the goals of ranking is to encourage students to select top universi-

ties in terms of education. From this perspective, international ranking that do not cover a large number of universities across countries cannot be used. The present ranking model is designed for a national scale and can, therefore, be used as a source of information for students. Educational indicators- especially those that measure outcomes- are a good benchmark for students and process indicators can also be used for this purpose. The results of these rankings were published and distributed in Persian in 2 separate booklets

Another important aspect of the existing ranking was their focus on outcome indicators. Outcome indicators are easier to assess (1, 4), while measuring educational outcomes would be very difficult or even impossible in some cases. Therefore, ranking systems might not pay much attention to educational outcomes (1). In contrast, the present model provides a more comprehensive view of the education system by combining process and outcome indicators.

Most ranking systems use complex indicators that are not easily understood by most students or even the scientific community. We tried to design simple and reliable indicators that can be easily used by students, faculty, and policymakers. The present indicators are representative of the main educational activities; therefore, the system administrators can use them to monitor and evaluate their functions. Although they are designed relevant to the educational processes within the UMS in Iran, they can be modified or replaced by other sets of indicators to fit the target setting. Our results showed stability over time for the indicators and maximum conceptual consistency with other indicators. The stability and consistency make them suitable for use.

Although universities in different countries have different structures, the approach used in the design of the present model may help adjust indicators according to the country's other needs.

There were high internal consistency and high correlations between indicators. The structure of factors was also consistent with the theoretical framework of the study.

Although high internal consistency and factorial structure are the 2 main concerns of ranking systems, the results of previous studies have shown that all ranking systems are not equally reliable and the fit of their factor structure is not comparable (18-20). In this study, high internal and longitudinal consistency and good structure fit was found for measured indicators. The main reason for these results was that all indicators were related to the university's educational performance. Usually, ranking systems and league tables assess different indicators related to the various functions of universities, and the results of previous factor analyses also showed that a small number of indicators explain a significant percentage of variance (18). The structure of universities within each country is similar, but there tends to be more differences across universities in different countries. Designing indicators based on the educational structure of each country can also help to design a better structure. Accordingly, Soh (18) argued that to avoid misinterpretation of the overall results, there is a need to reconceptualize university rankings to have a higher-order institutional excellence, comprising 2 lower-order measures based on academic excellence and managerial excellence.

The increase of the scores in most of indicators in the second round was also one of the main results of this study. The results of previous research have also shown that entering the competition would improve universities' scores. However, various studies looking at the effectiveness of ranking in the performance of universities has shown different results (1, 16). Shin and Toutkoushian (3) described the impacts of rankings on higher education systems, universities, and students, and also explained the side effects. They concluded that to have a positive impact, the current unified ranking systems should become multiple ranking systems to reflect different institutional missions, rater-centered systems should become customer-centered, and global ranking systems should become regional. In the present ranking system, we utilized a regional and customer-centered system that reflects the educational missions of medical universities. Performance improvement may be due to the choice of this approach.

One of the limitations of this research is its relative dependence on the educational context of Iran's UMS. This restriction reduces the possibility of transferring the indicators to other contexts. Nevertheless, its process-based nature makes it possible to translate it into various contexts. Another limitation is the difficulty of examining educational consequences and outcomes, which is a problem for most ranking systems. We used 2 main outcome variables, including postgraduate education entrance and internal consistency of students' scores in professional courses, both of which are direct consequences of the education system. These are short-term consequences as assessing the longer ones is much more difficult- especially in the context of medical disciplines. The lack of a link between educational practice and other university practices, such as research, is one of the constraints when analysing the position of universities. Research rankings of universities are also conducted in Iran, but they use indexes similar to the international rankings, and they pay less attention to process indicators.

One of the most important strengths of this ranking system was its 2-sided perspective. The process of indicator design was mutual and the universities had active involvement. The members of the design team were the faculty members of the UMS. The views of the university deputies were taken at several stages. They were asked about the importance, measurability, precision, and improvability of the indicators. Two rounds of data collections with participation of UMS indicated the practicality of indicators.

## Conclusion

Based on the results of both rounds of rankings, the best performance was observed in the educational endpoint indicators, which were the most objective indicators. The result also showed that universities were more successful in indicators that were more visible. The worst performance belonged to those universities that followed the National Scientific Roadmap because of its indicators having a more abstract nature. Although the National Scientific Roadmap has shown the desired strategies, it has not been as successful in showing the way to achieve these goals. Also, its concepts were not translated into tangible operational objectives and minimum enforcement was implemented to regulate university performance.

This ranking has improved the performance of universities. It has also helped with their efforts to improve processes. One of the achievements of this system was the alignment of processes based on the structure of the universities. The universities gave positive feedback after running both rounds. Given the positive results that followed, the Ministry of Health plans to launch the third round. There are awards for top universities and top universities in indices and domains also receive awards.

## Acknowledgments

### Conflict of Interests

The authors declare that they have no competing interests.

## References

1. Çakır MP, Acartürk C, Alaşehir O, Çilingir C. A comparative analysis of global and national university ranking systems. Scientometrics. 2015 Jun; 103(3):813-48.
2. Aguillo IF, Bar-Ilan J, Levene M, Ortega JL. Comparing university rankings. Scientometrics. 2010 Oct; 85(1):243-56.
3. Shin JC, Toutkoushian RK. The past, present, and future of university rankings. University Rankings: Springer; 2011. p. 1-16.
4. Dearden JA, Grewal R, Lilien GL. Framing the university ranking game: actors, motivations, and actions. Ethics Sci Environ Politics. 2014 Jun; 13:131-9.
5. Marginson S. Global university rankings: Implications in general and for Australia. J High Educ Pol Manag. 2008 Apr; 29(2):131-42.
6. Florian RV. Irreproducibility of the results of the Shanghai academic ranking of world universities. Scientometrics. 2007 Feb; 72(1):25-32.
7. Shin JC, Toutkoushian RK, Teichler U. University rankings:

Theoretical basis, methodology and impacts on global higher education. London: Springer. 2011.

8. Ioannidis JP, Patsopoulos NA, Kavvoura FK, Tatsioni A, Evangelou E, Kouri I, et al. International ranking systems for universities and institutions: a critical appraisal. BMC Med. 2007 Oct; 5(1):30.

9. Jeremic V, Bulajic M, Martic M, Radojicic Z. A fresh approach to evaluating the academic ranking of world universities. Scientometrics. 2011 Jun; 87(3):587-96.

10. Molinari JF, Molinari A. A new methodology for ranking scientific institutions. Scientometrics. 2008 Apr; 75(1):163-74.

11. Lopez-Valcarcel BG, Ortún V, Barber P, Harris JE, García B. Ranking Spain's Medical Schools by their performance in the national residency examination. Rev Clin Esp. 2013 Dec; 213(9):428-34.

12. Arora J, Trivedi K. Developing National Ranking for Indian Universities: Moving towards World Class Universities. Inflibnet. 2014 Jul; 21(2):20-7.

13. Alaşehir O, Çakır MP, Acartürk C, Baykal N, Akbulut U. URAP-TR: a national ranking for Turkish universities based on academic performance. Scientometrics. 2014 Oct; 101(1):159-78.

14. Dill DD, Soo M. Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. High Educ. 2005 Jun; 49(4):495-533.

15. Lukman R, Krajnc D, Glavič P. University ranking using research, educational and environmental indicators. J Clean Prod. 2010 May; 18(7):619-28.

16. Hazelkorn E, Gibson A. Another Year, Another Methodology: Are Rankings Telling Us Anything New? Int High Educ. 2016; (84):3-4.

17. Williams R, de Rassenfosse G, Jensen P, Marginson S. The determinants of quality national higher education systems. J High Educ Pol Manag. 2013 Nov; 35(6):599-611.

18. Soh K. What the Overall doesn't tell about world university rankings: examples from ARWU, QSWUR, and THEWUR in 2013. J High Educ Pol Manag. 2015 May; 37(3):295-307.

19. Steiner JE. World university rankings-A principal component analysis 2006 [Available from: https://arxiv.org/ftp/physics/papers/0605/0605252.pdf.

20. Aguillo IF, Bar-Ilan J, Levene M, Ortega JL. Comparing university rankings. Scientometrics. 2010; 85(1):243-56.