Development of a local model for measuring the work of surgeons

Sara Forootan1, Sakineh Hajebrahimii2, Ali Janati3, Behzad Najafi4, Mohammad Asghari-Jafarabadi4

1 Department of Health Policy and Management, Tabriz University of Medical Sciences School of Management and Medical Informatics, Tabriz, Iran
2 Research Center for Evidence Based Medicine (RCEBM), Tabriz University of Medical Sciences Faculty of Medicine, Tabriz, Iran
3 Department of Health Economics, Tabriz University of Medical Sciences School of Management and Medical Informatics, Tabriz, Iran
4 Center for the Development of Interdisciplinary Research in Islamic Sciences and Health Sciences, Tabriz University of Medical Sciences, Tabriz, Iran

ABSTRACT

Objective: The Relative Value Unit (RVU) is the main method of calculating surgeons’ reimbursements and a tool for measuring the work of surgeons. Existing evidence shows that the work Relative Value Unit (wRVU) does not accurately represent surgeon’s work. Therefore, the current study attempted to develop a local model to measure surgeons’ work.

Material and Methods: This study was conducted in two main phases of determining the metrics and model development using quantitative and qualitative approaches from December 2019 to April 2021. Literature review, focused group discussions, and interviews were used to collect data. Content analysis and Exploratory Data Analysis techniques were applied to analyze data.

Results: The findings demonstrated that patient’s conditions (age, severity of disease at referring time, and comorbidities), disease specifications (time, complexity, physical effort, and risk), and provider characteristic (surgeon’s willingness, imposed stress, and surgeon’s skill) were important by 17, 51, and 32%, respectively, in determining surgeons’ work.

Conclusion: Determining a fixed value for each procedure does not accurately estimate the amount of required surgeon’s work for any procedure. Many factors, such as the patient’s condition, surgeon’s characteristics, and disease specification affect surgeons’ work in the operation room. Proper measurement of the surgeon’s work is an important step towards establishing equity in payment in the health system.

Keywords: Relative value unit, work relative value unit, surgeon’s work, measurement

INTRODUCTION

In response to rising health care expenses in the 1980s, Medicare took a new reimbursement approach (1). This new approach was the consequence of studies by Hsiao et al. at Harvard, which was known as Resource-Based Relative Value Scale (RBRVS) (2). This reimbursement method was defined based on the resources spent for each service specified by CPT codes (3). RBRVS consists of three components, the most important of which is work Relative Value Unit (wRVU). The first reason for its importance is because it accounts for 52% of RBRVS and, the second one is that it is a tool that determines the amount of physicians’ work for a service, in a way, the reimbursement to the physician (4). According to Hsiao et al., the components of wRVU include the time required to perform a procedure, mental effort and judgment, the physician’s physical effort and technical skills, and stress (5). wRVU is now not only used as a reimbursement model but is also known as a criterion to measure a physician’s performance and productivity (6). Although RVUs are reviewed by RUC based on receiving suggestions and criticisms from physicians annually, there are still criticisms of RVUs in surgery (7). Previous research showed that the RVU-based payment system did not accurately represent the surgeon’s work. According to studies, there is a weak correlation between the surgeon’s work and its metrics (8-10). The complexity of the surgeries is not well taken into account, the patient’s need in some surgeries to follow-up is not considered in the RVU and, consequently, in the payment to surgeons (11). An imbalance between
RVU-based system and the amount of work a surgeon performs causes dissatisfaction and burnout in the surgeons' community, which may have irreparable consequences such as behavioral changes in surgeons (10,12). Due to the importance of proper reimbursement to surgeons as one of the most important elements in the health system and its effect on the quality of care, our research aimed at finding a solution for this challenging problem by focusing on two scopes, identifying metrics to measure Iranian surgeons' work and providing a relevant model.

**MATERIAL and METHODS**

This study was conducted in two main phases of determining the metrics and model development using quantitative and qualitative approaches from December 2019 to April 2021 in Iran.

**Literature Review**

A comprehensive literature review was conducted to identify metrics that measure surgeons' work in Google Scholar, PubMed, EMBASE, Scopus, Ovid Medline, Web of Science, Cochrane, ProQuest, Scientific Information Database (SID), and Magiran databases. Search keywords were “surgery”, “reimbursement mechanism” “surgeons’ work”, “surgery fee”, “physician fee”, “surgeon workload”, “compensation”, “relative value units”, “RBRVS”, “wRVU”, and a combination of these keywords, along with searching Persian databases. All searches were conducted without a time limitation. The World Health Organization (WHO) and CMS.Gov websites were also reviewed here. Selected articles were related to RVU and surgeons’ work, and duplicate articles were excluded from the study. Finally, metrics related to the surgeon’s work were identified after reviewing articles.

**Focused Group Discussion and Interview**

Since some of the metrics found in the literature review were not available and feasible in Iran, Focus Group Discussion (FGD) and interviews were conducted to identify metrics relevant to the country concept. The purpose of FGD is to increase the quality of data through group dynamics (13). Three FGD meetings were held with surgeons from eight specialties (Urology, Gynecology and obstetrics, Neurological, Ophthalmic, Orthopedic, Cardio-thoracic, Otorhinolaryngology, and General surgery).

Each meeting lasted about 120 min, with the attendance of a total of 30 surgeons. The coordinator asked questions during the meetings and tried to involve all participants in the discussions. The surgeons were asked questions about “In your opinion, which metrics should be considered to measure the surgeon's work?” and “In your point of view, what are the missing metrics in calculating current wRVU?” Finally, they were asked to introduce a surgeon as the representative of that specialty to connect the research team to other surgeons in that specialty for the next step of the study.

Semi-structured interviews were conducted with ten surgeons who were not able to attend the FGDs. The interviews continued until saturation, the average interview time was 45 min and were carried out in the surgeon’s office. The surgeons were selected purposefully to participate in this study (14). The inclusion criteria were surgeons interested in participating in the study, surgeons who earned by RVU payment method and were familiar with the RVU concept, and those who had an article or some research related to RVU and were experts in their specialty.

**Data Analysis**

**Qualitative Analyses**

This study was conducted with an inductive approach in the form of content analysis. Data analysis began concurrently with data collection aiming at using these analyses to help shape the next steps in data collection. Two researchers transcribed and analyzed FGD and interviews on paper. Coding was done in the margin. The main themes were identified, and the relationship between themes was recognized by the research team, followed by merging similar ones. The research team returned a summary of the notes to the participants, and they confirmed the accuracy of the data to increase the study accuracy and rigor (15,16).

**Metrics Selection**

The identified metrics of the comprehensive review, FGD, and interviews were integrated into a questionnaire consisting of a list of metrics to measure the surgeon’s work. The surgeons were asked to select the metrics that were important to measuring the surgeon’s work. Questionnaires were provided to surgeons in eight specialties through representatives. Finally, 91 questionnaires were returned from 100 distributed questionnaires.

**Model Development**

Next, related metrics were placed in a group, and the surgeons were asked to compare the metrics in pairs and score them based on similarities selected from the previous stage with a score of 50% or above. The similarity and dissimilarity between the two metrics were scored from 10 to 1. The groups had to be prioritized and weighted in the final step. A questionnaire was designed to compare the groups in pairs by two criteria of necessity and effectiveness. The questionnaires were given to the surgeons, who were asked to rank the groups based on the two mentioned criteria. Finally, 87 questionnaires were returned from 100 distributed questionnaires.

**Delphi Technique**

A Delphi questionnaire model was developed for approval after grouping and weighing the metrics. The final metrics were distributed to 100 specialists in the form of surgeons’ work mea-
measurement model. Experts were asked to score the metrics with a 9-point Likert scale based on three criteria of importance, simplicity and clarity, and feasibility. The median was used to calculate the score of each metric.

**Statistical Analysis**
For statistical analysis, the similarity of metrics was examined using the Multidimensional Scaling method. The Exploratory Data Analysis technique was applied to categorize the relevant indicators in a group by Stata (version 16). The Multiple Criteria Decision Making (MCDM) approach was used for weighting groups with Super Decision (version 3). The mean and median of the indicators were calculated by Excel (2013).

**Ethical Consideration**
The study is a part of a Ph.D. dissertation with the ethics code IR.TBZMED.REC.1397.960. Participation in this study was optional, all members participated with informed consent and were notified that their information would remain confidential and anonymous. The participants had the right to withdraw from the study at any time. They were allowed to record audio during FGD and interviews.

**RESULTS**

**Participant Profile**
Forty participants attended the FGD and interviews. Age range was between 36 and 68 years, and 80 and 20% of participants were men and women, respectively. Their work experience was between 6 and 35 years. Participants were specialized in Urology (4), Gynecology and obstetrics (3), Neurological (3), Ophthalmic (3), Orthopedic (4), Cardiothoracic (3), Otohinolaryngology (4), General surgery (4), and healthcare management (2).

**Metrics Selection**
A comprehensive literature review was done by related keywords. A primary review resulted in a total of 105 articles, 14 of which were eliminated for duplication, and 37 studies were excluded because of lacking related information. Finally, 54 articles were included in the study. Then, 19 metrics were derived through literature review (Table 1), and 21 metrics were obtained from FGD and interviews. Eleven metrics were removed since they were noted in both the literature review and the qualitative data. Finally, 29 metrics were selected and made available in the form of a questionnaire. The surgeons were asked to select the effective ones in the wRVU calculation. The average score was calculated for each metric. According to the research team, selected metrics had a grade of 50% and above (meaning that at least 50% of surgeons selected this metric as a significant item to calculate their work), and metrics with scores below 50% were excluded from the study, ultimately choosing 12 metrics. Selected metrics were patient age, the severity of the disease, operation duration, risk, the complexity of the surgery, imposed stress on a surgeon during surgery, surgeons’

<table>
<thead>
<tr>
<th>Number</th>
<th>The metric</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Operation Time</td>
<td>(7,9,17-20)</td>
</tr>
<tr>
<td>2</td>
<td>Mental effort</td>
<td>(21-23)</td>
</tr>
<tr>
<td>3</td>
<td>Physical effort</td>
<td>(20,24,25)</td>
</tr>
<tr>
<td>4</td>
<td>Psychological stress</td>
<td>(20,26)</td>
</tr>
<tr>
<td>5</td>
<td>Technology</td>
<td>(10,27)</td>
</tr>
<tr>
<td>6</td>
<td>Patient characteristic</td>
<td>(8,28-33)</td>
</tr>
<tr>
<td>7</td>
<td>Mortality risk</td>
<td>(20,29,4)</td>
</tr>
<tr>
<td>8</td>
<td>Physician willingness to provide a service</td>
<td>(35)</td>
</tr>
<tr>
<td>9</td>
<td>Complexity</td>
<td>(7,9,19,31,36,37)</td>
</tr>
<tr>
<td>10</td>
<td>Severity and Emergency operation</td>
<td>(38-41)</td>
</tr>
<tr>
<td>11</td>
<td>Quality of service and outcome</td>
<td>(10,37,39,42-47)</td>
</tr>
<tr>
<td>12</td>
<td>Number of attending staff at OR</td>
<td>(18,48)</td>
</tr>
<tr>
<td>13</td>
<td>Academic rank of surgeon</td>
<td>(8,17,49-51)</td>
</tr>
<tr>
<td>14</td>
<td>Skill and surgeon’s experience</td>
<td>(8,40,45,46,48)</td>
</tr>
<tr>
<td>15</td>
<td>Surgeon’s efficiency</td>
<td>(45,52,53)</td>
</tr>
<tr>
<td>16</td>
<td>Pre and post operation time</td>
<td>(8,25,36,48,54,55)</td>
</tr>
<tr>
<td>17</td>
<td>The leadership and coordination role of surgeon in OR</td>
<td>(43,46)</td>
</tr>
<tr>
<td>18</td>
<td>Type of hospital (teaching/community)</td>
<td>(50,56)</td>
</tr>
<tr>
<td>19</td>
<td>Demand for a service</td>
<td>(21,24,35,39,47)</td>
</tr>
</tbody>
</table>
willingness to operate, skill, physical effort during surgery, comorbidities, pre-operation time, and post-operation time. Complexity gained the highest average score of 93%.

**Grouping Metrics**

Multi-Dimensional Scaling method was used to classify similar metrics in a group, and the metrics were compared and scored in pairs. Minimum and maximum scores were 1 and 10, respectively. The findings of this step were in the form of a matrix in which the horizontal and vertical axes consisted of 12 metrics. The comparison of the score of each metric in the horizontal axis with itself in the vertical axis was zero. An entry of the matrix showed the geometric mean of the similarity score of the two indicators based on the respondents’ opinions. The findings were analyzed by Exploratory Data Analysis. Patients’ age, the severity of the disease, and comorbidities were placed in a group as the patient’s condition. The metrics of operation duration, risk, complexity, physical effort, and pre-and post-operative times were categorized in a group as the disease specification, and the metrics of imposed stress, surgeon willingness, and skill were categorized as surgeon’s characteristics.

**Prioritizing the Groups**

The AHP method was used to determine the importance and weight of each group of metrics. The three groups containing the indicators were compared to each other in pairs and scored based on two criteria of necessity and effectiveness. The scoring scale ranged from 1 to 9. Both the necessity and effectiveness criteria were considered equal with equal weight (0.5). The obtained data from the questionnaires were entered into Super decision 3 software in the form of a weighted average, which was the result of experts’ opinions. After data analysis, the patient’s condition, disease specification, and provider characteristics were weighted as 17, 51, and 32%, respectively. In this study, inconsistency rate was 0.05. A summary of the grouping is given in Table 2.

**Delphi**

The model was confirmed by the Delphi technique. According to the research team, approved metrics had a median score of 7 or higher. In the present study, minimum and maximum scores for the metrics were 8 and 9, respectively. This model was confirmed by a single step of Delphi. Finally, the expert panel of nine surgeons was held to confirm the model. According to the decision of the experts present in the panel, the length of operation, pre-operation, and post-operation times were merged and named in the form of a single metric as time. A summary of the study process is shown in Figure 1.

**DISCUSSION**

RVUs were developed to reduce healthcare expenditures and Medicare costs (17). wRVU, which measures surgeon’s work for a particular service, has been gradually considered as an important indicator of productivity, performance, and eventually payment for surgeons. In the last decade, more attention has been paid to wRVU and its metrics due to the importance of equity in payment, the proportion of a surgeon’s work to earn, and the desire of surgeons to perform certain surgeries, which Hsiao et al. did not expect to be one of the most important challenges to the health system in the coming decades. Proper measurement of the surgeon’s work is a prerequisite for a proportionate payment system. After three decades, this study provided a native model for measuring surgeons’ work. The findings of our study reveal that measuring surgeons’ work solely based on metrics, such as operation duration, risk, physical effort, and mental effort, does not accurately reflect surgeon efforts in the operation room (OR). What happens in the OR is more than that. These results go beyond previous reports, showing that RVUs do not accurately measure the time and effort of procedures across many subspecialties (18,19).

According to the results, several factors influence surgeon’s effort in OR, such as the patient’s age, disease severity at referral, preoperative consultation time, postoperative care time, operation duration, surgical risk and complexity, the stress imposed on the surgeon during the operation, surgeons’ willingness to operate, skills, physical effort, and comorbidities.

As commented by the participating surgeons, severity of the disease at the referring time, patient’s conditions such as hypertension or diabetes during the operation, and whether the patient is an elderly man in the last years of his life, a child with a high life expectancy, or a young man in his 25s do not make a difference in the patient treatment by the surgeon in the OR, but the stress transferred to the surgeon in OR is far from the payers’ view. Schwartz et al. state that RVU does not distinguish extra work required by an emergent patient (20).

Due to the change in people’s lifestyle, comorbidities are more common than 30 years ago (21), which not only make surgery
more stressful for both the surgeon and the patient but also can lead to postoperative complications. Therefore, these patients need more attention and effort in the OR. Based on this study, patients' conditions have a 17% effect on the amount of physician's work in the OR. As stated above, the age of the patients should be taken into account in determining the relative value (22-24). The findings are directly in line with previous findings, and similar studies have emphasized paying attention to the patient's characteristics in determining the wRVU, which is neglected in the current RVU system (8,25,26).

In order to perform an operation, Hsiao considers time as an important factor in calculating the work of the physician (27). However, calculating the surgeon's work only based on the
length of a surgical operation causes bias. It may be necessary for a patient to consult a surgeon before surgery, or to be followed up with a surgeon for a long time after surgery in some cases (9). In contrast to measuring operation times, this study suggests considering pre- and post-operative care times as well. For each surgery, the term time refers to the time required for pre-operative consultation, the length of the operation, and the postoperative care required. This is consistent with that found in previous studies surgeons who spend time on such affairs as consulting, operative planning, and committee work, for which no payment is made (28,29). A study by Shah et al. has shown poor correlation between RVUs and operative time for a variety of high-volume surgical procedures (8).

From the results, it is clear that besides time, complexity is another metric to measure a surgeon’s work. Complicated surgeries take more time and effort and impose more stress on surgeons, therefore, negatively affect their willingness to perform such procedures. In this regard, these findings are consistent with research showing that complexity in operation may need more attention, time, and effort, and, therefore, would be considered in calculating the surgeon’s work (30-33).

It is important to highlight the fact that the surgeon’s characteristics as the service provider are effective in their work. The present study confirmed the findings about the amount of stress that a physician experiences during surgery, the surgeon’s skill during that operation, and the surgeon’s willingness to perform that operation affect wRVU by 32%. Therefore, it seems reasonable to measure a surgeon’s work considering this finding.

An operation performed by a more skilled surgeon often results in fewer complications for the patient and a shorter surgical duration. Consequently, it affects efficiency and cost of the health system. Previous studies have also noted the weakness of the RVU system in not considering the surgeon’s skill (8,34). Also, patients may have difficulty accessing certain specialties if the surgeon is unwilling to perform certain operations (35,36).

Other studies have found that metrics, such as the quality of care, patient satisfaction, and the technology used in the procedure, would be considered to measure the surgeon’s work, which are not mentioned in our study. This can be attributed to differences in payment structures, inadequate and unreliable data (physical and electronic) of surgical complications, and medical errors. Because of this potential limitation, it was impossible to measure such metrics (10,37-40).

Nonetheless, we believe that determining a certain value for each procedure does not accurately estimate the amount of required work for a procedure because, in addition to the disease specification, it is also affected by many factors such as the patient’s condition, the surgeon’s skill, and the provider’s tendency. Therefore, we suggest that a range of values with a minimum and a maximum should be considered instead of a fixed wRVU in the RVU schedule for a procedure. The surgeon’s skill and willingness, operation complexity, and the patient’s condition will determine the value of minimum or maximum.

As discussed above, applying this model to determine wRVU causes similar RVU of a procedure to vary in different situations. In addition to ensuring fair payment for surgeons, it would also ensure that patients have access to the required procedures.

**CONCLUSION**

Due to wRVU’s direct effect on payment, measuring the surgeon’s work is one of the most challenging issues related on one hand to the surgeon’s satisfaction and the health system expenditures on the other. A rational and accurate measurement of the surgeon’s work is an important aspect of establishing equity within the health system which is the initial mission of health systems. Performing procedures have now changed considerably, therefore, the need to pay attention and review the metrics of work measurement is felt more than before. In addition to disease specification, the present study emphasizes the need for paying attention to the patient’s condition and the surgeon/provider’s characteristics for work measurement.

---

**Ethics Committee Approval:** This study approval was obtained from Tabriz University of Medicine Sciences (Decision no: IR.TBZMED.REC.1397.960).

**Peer-review:** Externally peer-reviewed.


**Conflict of Interest:** The authors have no conflicts of interest to declare.

**Financial Disclosure:** The authors declared that this study has received no financial support.

**REFERENCES**

3. Chiu RG, Siddiqui N, Nunna RS, Patel S, Rosinski CI, Chaker AN, et al. Association of neurosurgical work relative value units with objective markers of operative complexity. World Neurosurg 2021; 146: e194-e204. [CrossRef]


34. Brent RJ, Patel M. Physician rewards for different kinds of service: the RBRVS versus the CPR system. Plast Reconstr Surg 1997; 100(1): 51-7. [CrossRef]


Cerrahların iş yükünü ölçmede yerel bir modelin geliştirilmesi

Sara Forootan, Sakineh Hajebrahimi, Ali Janati, Behzad Najafi, Mohammad Asghari-Jafarabadi

1 Tebriz Tıp Bilimleri Üniversitesi Yönetim ve Tıp Bilimi Fakültesi, Sağlık Politikası ve Yönetimi Anabilim Dalı, Tebriz, İran
2 Kanîta Dayalı Tıp Araştırma Merkezi (RCEBM), Tebriz Tıp Bilimleri Üniversitesi Tıp Fakültesi Tebriz, İran
3 Tebriz Tıp Bilimleri Üniversitesi Yönetim ve Tıp Bilimi Okulu, Sağlık Ekonomisi Anabilim Dalı, Tebriz, İran
4 İslam İlimleri ve Sağlık Bilimleri Disiplinlerarası Geliştirme Merkezi, Tebriz Tıp Bilimleri Üniversitesi, Tebriz, İran

ÖZET

 Giriş ve Amaç: Bağıl Değer Birimi (BDB), cerrahların ödemelerini hesaplamada ve yine cerrahların iş yükünü ölçmede kullanılan ana yöntemdir. Mevcut veriler, iş yükü Bağıl Değer Biriminin (iyBDB) bir cerrahın iş yükünü kesin olarak temsil etmediğini göstermektedir. Bu sebeple, bu çalışmada cerrahların iş yükünü ölçmede kullanılabilecek yerel bir modelin geliştirilmesi amaçlandı.


 Bulgular: Cerrahların iş yükünü belirlemekte hastalanın durumu (yaş, başvuru esnasında hastalığın ciddiyeti ve komorbiditeler), hastalığa ilişkin ayrıntılar (zaman, karmaşalık, fiziksel efort ve risk) ve sağlık uzmanının özellikleri (cerrahın istekliliği, maruz calınan stres ve cerrahın becerisi) sırasıyla %17, %51 ve %32'lik oranlara önemli bir adımdır.


 Anahtar Kelimeler: Bağıl değer birimi, iş yükü bağıl değer birimi, cerrahların iş yükü, ölçme

DOI: 10.47717/turkjsurg.2021.5473