ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/medsci.2022-88728

Assessment of Assistant Physicians' Competencies for Making Accurate Diagnosis with Agreement Coefficients: A Methodological Study

Asistan Hekimlerin Doğru Tanı Koyabilme Yeterliliklerinin Uyum Katsayıları ile Değerlendirilmesi: Metodolojik Bir Çalışma

İsmet DOĞAN[a], Nurhan DOĞAN[a]

[a]Department of Biostatistics and Medical Informatics, Afyonkarahisar Health Sciences University Faculty of Medicine, Afyonkarahisar, Türkiye

ABSTRACT Objective: The objective of this study is to introduce some statistical methods, which are thought to be used in evaluating the ability of resident physicians to make the correct diagnosis, with a sample application. **Material and Methods:** The most common way to show the accuracy of a decision maker's decisions relative to the decisions to be referenced is to construct an error matrix. In the ideal case, all off-diagonal elements of the error matrix would have a value of zero, indicating that no decision of the decision maker was wrong with respect to the reference. The calculations of the methods in the study are shown on a sample error matrix. The data in the error matrix are randomly derived data. In the evaluation of the error matrix, the agreement coefficients suggested by Bishop et al., Helldén and Short were used. **Results:** When the decisions of the decision maker (assistant physician) and the reference (specialist physician) decisions are similar, the variance of the error matrix decreases, the confidence intervals narrow, and the value of the coefficients of agreement both in the categorical and in general increases. **Conclusion:** Objective assessment techniques rather than subjective assessment are important for the standardization and quality of education. According to the findings, it was concluded that the agreement coefficients for proficiency in specialization education are usable and the quality of education will increase if they are used together with other evaluation methods.

Keywords: Specialization education; categorical agreement; competence; making a diagnosis

ÖZET Amaç: Bu çalışmanın amacı, asistan hekimlerin doğru tanı koyma yeterliliklerinin değerlendirilmesinde kullanılabileceği düşünülen bazı istatistiksel yöntemleri örnek bir uygulama ile tanıtmaktır. **Gereç ve Yöntemler:** Bir karar vericinin vermiş olduğu kararların referans kararlara göre doğruluğunu göstermenin en yaygın yolu bir hata matrisi oluşturmaktır. İdeal durumda, hata matrisinin tüm köşegen dışı elemanlarının değeri sıfır olur ve bu da referansa göre karar vericinin hiçbir kararının yanlış olmadığını gösterir. Çalışmada yer alan yöntemlerin hesaplamaları örnek bir hata matrisi üzerinde gösterilmiştir. Hata matrisinde yer alan veriler rastgele türetilmiş verilerdir. Hata matrisinin değerlendirilmesinde Bishop ve ark., Helldén ve Short tarafından önerilen uyum katsayıları kullanılmıştır. **Bulgular:** Karar verici (yardımcı hekim) kararları ile referans (uzman hekim) kararları birbirine benzedikçe hata matrisine ait varyans küçülmekte, güven aralıkları daralmakta, gerek kategori gerekse genel olarak uyum katsayılarının değeri büyümektedir. **Sonuç:** Öznel değerlendirmeden ziyade nesnel değerlendirme teknikleri eğitimin standardizasyonu ve kalitesi için önemlidir. Elde edilen bulgulara göre uzmanlık eğitiminde yeterlilik için uyum katsayılarının kullanılabilir olduğu ve diğer değerlendirme yöntemleri ile birlikte kullanılması hâlinde eğitimin kalitesinin yükseleceği sonucuna ulaşılmıştır.

Anahtar Kelimeler: Uzmanlık eğitimi; kategorik uyum; yeterlilik; tanı koyma

Correspondence: Nurhan DOĞAN
Department of Biostatistics and Medical Informatics,
Afyonkarahisar Health Sciences University Faculty of Medicine,
Afyonkarahisar, Türkiye
E-mail: nurhandogan@hotmail.com

Peer review under responsibility of Turkiye Klinikleri Journal of Medical Sciences.

Received: 04 Feb 2022 ***Received in revised form:*** 10 Oct 2022 ***Accepted:*** 10 Oct 2022 ***Available online:*** 02 Nov 2022

Assistant physicians (residents) are an important element of healthcare delivery. One of the more challenging performance issues that assistant physicians must be mastered is the ability to make an accurate diagnosis. Making a diagnosis is a complex cognitive proess that requires an assistant physician to possess of medical and practical knowledge that they can apply to often ill-defined clinical problems.[1] Competency-based medical education has emerged as a priority topic for medical education planners in the early 21st century.[2] Although the concept of competence is focused primarily on the trainees (residents), the need to clearly define the knowledge and skills needed by the trainers today is also considered important.[3] A significant number of medical adverse events occur due to errors in making a diagnosis. Hence, program evaluation is an essential responsibility for anyone overseeing a medical education program of assistant physicians.[4] The quantity and quality issues of the physicians who provide health services in the institutions still prevail today. Problems in the basic and continuous education of almost all physicians serving in a variety of healthcare facilities are the main reasons for the inadequacy in health service delivery. Logically, the provision of a quality health service depends on the training of physicians who take part in health care services. Medical education is an expensive, labor-intensive, and difficult process that needs qualified and devoted manpower in today's conditions.[5] Medical education in ancient ages and clinical training in the modern times is a process that has been still carried out within the framework of the "master-apprentice" relationship. Those who are trained in this model are like a copy of those who trained them. However, the role of the masters in deciding whether they are adequate for this training and when they become competent is controversial. Today, a competency-based education approach is taken into consideration while creating the goals and objectives of the specialty education and the programs to be followed to achieve these goals. Competence-based traning is an educational approach that ensures the training of physicians who are capable and competent in knowledge, skills, values, and behavior. In this approach, a model of complete learning by doing and practising is adopted.[6] It is stated that the term competence, which forms the basis of clinical evaluation, *per se* is difficult to define and measure, and it is claimed to be an ambiguous concept that is defined differently by different people. Its relation to other concepts such as talent, performance, and expertise is also unclear. Although it is generally accepted that clinical competence is crucial for making safe and effective decisions regarding diagnosis and treatment, there is no standard definition of the concept. The three ways of understanding competence are expressed as combining a set of general qualities such as task-related skills, general qualifications required for effective performance, and appropriate knowledge, skills, and attitudes for professional practice.[7] By Epstein and Hundert, professional competence in medicine is described as the usual and logical use of communication, knowledge, technical skills, clinical reasoning, emotion, value, and thought in daily practice for the benefit of the individual and society served. It is also stated that the basic clinical skills are built on the foundation of scientific knowledge and moral development.[8] Quality standards of specialty education in medicine should be developed and implemented for both the training programs and the educators. Three important factors determined to be successful in the quality improvement initiatives of specialty education programs in medicine are listed as purpose, data-based changes, and continuous improvement process.[9] The understanding that the best way to evaluate the quality of training provided in specialty training is to evaluate the quality of its products has become dominant. For this reason, it is stated that the standardization process will encounter various resistance and will be painful. Basing this process on the measurements taken from the assistants' working in educational institutions, namely concrete data, is important in terms of objectively seeing what the items need to be corrected and thus avoiding the vicious discussions caused by subjective evaluations.[10] At the beginning of the education process, the realization of the desired behaviors is only expected. Although education is defined as the process of creating the desired behavioral change in an individual through internal motivation in his own life the desired behaviors may not always be fulfilled at the end of the education process. Sometimes there are unwanted

or desirable but not sufficiently acquired behaviors. Identifying these inadequacies and undesirable behaviors in the education system, revealing the root causes of these deficiencies, that is, monitoring and controlling the functioning of the system, and ensuring its improvement is only possible through measurement and evaluation.[11] The approach to specialty education in medicine is rather oriented towards the product and the expected output. Learning outcomes in output-oriented education are determined; the content of the training program and how it will be organized, the training strategies to be applied, the training methods, the evaluation methods, and the training environment are handled within the framework of the determined learning outcomes. Measurement and evaluation will become one of the most important issues in medical specialty education in the future.[12] In the literature, there are methods related to the measurement and evaluation of medical education, expressed as objective formal clinical exams, mini clinical evaluation exams, scenario suitability tests, portfolio, and 360-degree evaluations.[13] Different evaluation areas and measurement-evaluation methods suitable for these areas are collected under three headings;[14]

- Performance and work-based assessment (clinical causation and decision making, management, organization, professionalism, individual development, etc.),
- Competence, clinical representation, clinical and practical assessment (clinical skills, practical applications, communication skills, knowledge management, clinical processes, diagnosis, treatment, follow-up processes, and protection),
- Information-oriented measurement and evaluation (information on the field of medicine).

In the specialty education in medicine, it is necessary to acquire the equipment that will enable the assistant physician to cope with the cases he will encounter in his professional life. It is thought that an efficient health service can be provided with this equipment.[15] Especially in the field of health care, the things such as communication skills of physicians with patients and patient relatives, behavior by moral and legal rules, etc. are qualifications that should be acquired and evaluated during education. Patients expect physicians to be competent in diagnosing, planning management, and implementing practical procedures, and while doing so, they expect them to behave in a reasonable and attentive, and humane attitude.[16] However, final decisions are important in output-oriented education, and the competence to be gained in the specialty training is to gain the ability to make the correct diagnosis and determine the correct treatment method. The objectives of the specialty training required by the twenty-first century and the programs to be followed to achieve these goals are discussed extensively and the qualification criteria differ according to the areas of specialty. Therefore, the need for techniques to evaluate the accuracy of the information obtained during residency training cannot be underestimated. It is impossible to improve this training without the methods of measuring and comparing the accuracy and consistency achieved using different criteria. Performance evaluation is the process of determining how successful the assistant physician is in his/her job using various methods, giving him feedback, and creating a development plan that will strengthen his weaknesses as well as his strengths. In this type of evaluation, the assistant physician does not give simple answers as in paper and pencil tests, he creates a product and this product includes decisions regarding diagnosis and treatment to a significant extent. Performance evaluation is a valid and reliable method because it allows for multiple interviews, multiple evaluators, and evaluation at different times and different settings.[17] Performance-based assessment methods are recommended to be used especially in the evaluation of applied areas.[13] The purpose of this article is to introduce some statistical methods that are thought to be used in the assessment of the competencies of assistant physicians. It is not within the scope of this article to present all the theoretical and practical details of the methods described in the study. However, with this study, it is expected that even users with a little or no knowledge of statistical methods will see the usability of these methods. In this study, two different methods are introduced. The first method is concerned with calculating the overall amount of agreement between two evaluators, taking into account all decisions made at

the same time. The method is also a procedure that allows testing whether two different agreement values are statistically significant. The second method allows two decision-makers to calculate the amount of agreement between them for each specific decision option.

MATERIAL AND METHODS

The most common way to show the accuracy of decisions made by a decision-maker relative to the decisions to be referenced is to create an error matrix table. The classical error matrix representation is given in Table 1 for the case where n different decision categories (diagnosis, treatment method, etc.) are indicated by K_1, K_2,...K_n.

An error matrix provides many opportunities to assess candidates' reasoning and ongoing competence. This notation for conceptualizing the problem is important because it provides a visual overview of decisions and highlights two types of errors specific to decisions. The errors stated here are the error of defining a decision that is not included in the ıth decision category in the th decision category, and the reverse identification error that occurs by placing a decision that should be included in the ıth decision category in another category. In an error matrix, the columns are generally the reference that is assumed to be correct (decisions made by the specialist physician, laboratory results, etc.), while the lines show the decisions made by the assistant physician. The error matrix allows the assistant physician to determine his performance for individual categories as well as for his overall decisions. Ideally, the value of all non-diagonal elements of the error matrix would be zero, indicating that no decision of the assistant physician was wrong. Once the error matrix has been created, a very simple procedure can be used to determine overall accuracy. Since the values on the main diagonal represent decisions made correctly, these values are summed and divided by the total number of decisions classified. This number is the overall performance accuracy of an error matrix and is the most common use of the error matrix in accuracy evaluation. It has been suggested to use Equation 1 as the general agreement measure in calculating the agreement between two evaluators for each error matrix.[18]

$$\widehat{K} = \frac{N * \sum_{i=1}^{n} X_{ii} - \sum_{i=1}^{n} (X_{i+} * X_{+i})}{N^2 - \sum_{i=1}^{n} (X_{i+} * X_{+i})} \quad 1$$

The agreement value $(\hat{K})$, indicates how compatible the reference decisions are with the decisions of the assistant physicians'. It may also be desirable to evaluate the compatibility of two assistant physicians with each other. The value of the test statistic is calculated using Equation 2.[19,20]

$$\frac{\widehat{K}_1 - \widehat{K}_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \sim Z \quad 2$$

$(\hat{K}_1)$: the agreement value of the first assistant physician',

TABLE 1: The representation of classical error matrix.

		Specialist physician/Reference						
		K_1	K_2	K_3	...	K_{n-1}	K_n	Total
Assistant physician/ Decision maker	K_1	X_{11}						X_{1+}
	K_2		X_{22}					X_{2+}
	K_3			X_{33}				X_{3+}
	⋮				...			⋮
	$K_{(n-1)}$					$X_{(n-1)(n-1)}$		$X_{(n-1)+}$
	K_n						X_{nn}	X_{n+}
	Total	X_{+1}	X_{+2}	X_{+3}	...	$X_{+(n-1)}$	X_{+n}	N

X_{ii}: The number of patients diagnosed by the specialist and assistant physician for the ith disease category; X_{i+}: ith row total; X_{+i}: ith column total; K_i: ith category; N: Total number of observations.

($\hat{K}_2$): the agreement value of the second assistant physician',

The normal distribution of the test statistic given by Equation 2 means that the values in the standard normal distribution table prepared for different confidence levels can be used to determine whether the two matrices are significantly different. For example, if the value of the test statistic at 95% confidence level is outside the range of $-1.96 \leq Z \leq 1.96$ it is concluded that there is a significant difference between the assistant physicians' agreement. The variance of $\hat{K}$ is calculated using Equation 3.

$$\sigma_x^2(\hat{K}) = \frac{1}{N}\left\{\frac{\vartheta_1(1-\vartheta_1)}{(1-\vartheta_2)^2} + \frac{2(1-\vartheta_1)(2\vartheta_1\vartheta_2-\vartheta_3)}{(1-\vartheta_2)^3} + \frac{(1-\vartheta_1)^2(\vartheta_4-4\vartheta_2^2)}{(1-\vartheta_2)^4}\right\} \quad 3$$

$$\vartheta_1 = \sum_{i=1}^{n}\frac{X_{ii}}{N}$$

$$\vartheta_2 = \sum_{i=1}^{n}\frac{X_{i+} * X_{+i}}{N^2}$$

$$\vartheta_3 = \sum_{i=1}^{n}\frac{X_{ii}}{N}\left(\frac{X_{i+}}{N}+\frac{X_{+i}}{N}\right)$$

$$\vartheta_4 = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{X_{ij}}{N}\left(\frac{X_{j+}}{N}+\frac{X_{+i}}{N}\right)^2$$

The confidence interval in which the real value of $\hat{K}$ is found is calculated using Equation 4.[18]

$$\hat{K} - Z_{1-(\alpha/2)} * \sigma_\infty(\hat{K}) \leq \hat{K} \leq \hat{K} + Z_{1-(\alpha/2)} * \sigma_\infty(\hat{K}) \quad 4$$

The agreement value given by Equation 1 is calculated by taking into account all the information obtained from decision-makers. However, by making use of this information, it may be requested to determine the agreement between decision-makers for a certain category (depression, personality disorder, schizophrenia, neurosis, etc.). Calculating agreement for each category separately is important because the reference and the assistant physician' provide detailed information on which categories are the least or the most compatible. Calculation of categorical agreement will guide decision-makers to reconsider their decisions regarding the category they are least compatible with, rather than all of their decisions, especially in cases where overall agreement is low. It is suggested to use the formulas given in Equation 5, Equation 6, and Equation 7 in calculating the category agreement.[18,21,22]

$$\hat{\kappa}_B = \frac{NX_{ii} - X_{i+}X_{+i}}{NX_{i+} - X_{i+}X_{+i}} \quad 5$$

$$MA_H = \frac{2 * X_{ii}}{(X_{i+}+X_{+i})} \quad 6$$

$$MA_S = \frac{X_{ii}}{(X_{i+}+X_{+i}-X_{ii})} \quad 7$$

The calculations of the methods included in the study are shown using the data given in Table 2.

The data are completely randomly derived data showing the distribution of diagnoses considered made by two assistant physicians and a specialist physician (reference).

RESULTS

Considering the information given in Table 2, the accuracy and general agreement values between the reference decisions and the residents' decisions are 0.72, 0.63, and 0.77, 0.69, respectively. Category agreement values are given in Table 3.

As can be seen from the table, the agreement coefficients differ for each category. As the numerical values outside the main diagonal increase, the agreement between the decisions of the assistant physician and the specialist physician decreases. It would be appropriate to concentrate the training on the category in which the value of the agreement coefficient is low. The variances of each error matrix are $\sigma^2_\infty(\hat{K}_1)=0.0035$, $\sigma^2_\infty(\hat{K}_2)=0.0031$ respectively. Since the decrease in the variance value indicates that the value of the agreement coefficient increases, it is desired to take a value as close to zero as possible. The 95% confidence intervals obtained for the *($\hat{K}$)* values are $0.51 \leq \hat{K}_1 \leq 0.75$, $0.58 \leq \hat{K}_2 \leq 0.80$. When all categories are considered at the same time, the limit values of the confidence interval will increase as the general agreement value increases. It is said that there is no difference between the two assistant physicians' matching the reference since it is calculated as $z = -0.8166$.

DISCUSSION

Assessment of competence is important to ensure that assistant physicians' are making safe and effective

TABLE 2: Derived sample data for classical error matrix.

		Specialist physician diagnosis				
		Depression	**Personality disorder**	**Schizophrenia**	**Neurosis**	**Total**
Assistant Physician 1 Diagnosis	Depression	16	1	1	2	20
	Personality disorder	1	16	2	1	20
	Schizophrenia	4	2	17	0	23
	Neurosis	6	7	1	23	37
	Total	27	26	21	26	100
Assistant Physician 2 Diagnosis	Depression	20	3	1	1	25
	Personality disorder	1	18	2	3	24
	Schizophrenia	4	2	17	0	23
	Neurosis	2	3	1	22	28
	Total	27	26	21	26	100

TABLE 3: Category agreement values for assistant physicians.

Decision maker	*Category*	K_B *Bishop et al.*[18]	MA_H *Helldén*[21]	MA_S *Short*[22]
Assistant Physician 1	Depression	0.73	0.68	0.52
	Personality disorder	0.73	0.70	0.53
	Schizophrenia	0.67	0.77	0.63
	Neurosis	0.49	0.73	0.58
Assistant Physician 2	Depression	0.73	0.77	0.63
	Personality disorder	0.66	0.72	0.56
	Schizophrenia	0.67	0.77	0.63
	Neurosis	0.71	0.81	0.69

decisions in diagnosis and treatment, although competence is difficult to define and measure. Reliability and validity are essential for the assessment of competence, but it is claimed that reliability and validity are rarely taken into account in the clinical evaluation process. There is also a controversial issue as to what level of competence can be considered inadequate and therefore a decision-maker. Considering a series of decisions made by a person, if 90% is sufficient, should it be considered sufficient in practice, or should it be 100% sufficient? Puts the subject of competence open to discussion. Competence assessment is complex due to the lack of standardized methods. The literature includes reports of many strategies used in assessing competence, but the evidence supporting their effectiveness is often lacking. The quantitative methods used for evaluation ensure that the decisions made are expressed in numbers, thereby making the decisions more sensitive, obtaining objective evaluations, and standardizing the evaluations. The accurate assessment methods described in the study use discrete multivariate analysis techniques. Classification data is discrete because patients may or may not fall into a particular disease category. Thus, a patient can be classified as paranoia or schizophrenia, but not as half paranoia and half schizophrenia. The methods mentioned in the article are simple and understandable methods that show how the amount of agreement regarding the decisions of the evaluators can be measured using error matrices. After the matrices are created, the evaluation can be performed with a simple Excel application. The high rate of accuracy should not mean that the agreement value will also be high. As can be understood from the error matrices, the large diagonal value in the cell where the same categories intersect, among the diagonal elements used in determining the amount of agreement, indicates that the amount of adaptation to

be obtained for the category in question will be high, while the small amount of agreement for this category will be low. The index given in Equation 5 does not guarantee that the agreement remains between 0 (complete incompatibility) and 1 (exact agreement), but can take negative values. The index given in Equation 7, on the other hand, cannot reflect the high harmony between the evaluators numerically as much as the index given in Equation 6. Therefore, the index suggested by Helldén should be preferred to other indices. It is stated that the MA_H index has no probability or mathematical basis, and is an index put forward by Helldén which is completely logical.[23] Difficulties in measuring proficiency in residency training led researchers to search for different methods for measurement and evaluation. In this context, it is considered appropriate to use the methods included in the study as a measure of competence at the level of accurate diagnosis. Considering the common problems in qualification assessments, it can be said that the proposed methods are reliable and valid enough for use in the actual work environment as a measure of performance on the job. By using the proposed methods as an evaluation measure, it is possible in a relatively short time to draw reliable and valid inferences about the assistant physicians' ability to make an accurate diagnosis using a sample of patients. Also, the proposed methods appear to be generic enough to be used in many clinical applications and are not limited to specific professional roles. Moreover, the proposed methods can be a useful aid to the traditional performance evaluation or annual review process in a formative or summative sense as a self-assessment or peer review tool for clinical use. The proposed methods can be used as a part of a general continuous quality improvement program in residency training, in other words as a quality assurance method.

CONCLUSION

As a result, it was concluded that the agreement coefficients are usable for competencies of assistant physicians making an accurate diagnosis in specialization education and the quality of education will increase if they are used together with other assessment methods. By using the agreement coefficients suggested in the study, the diagnostic competencies of assistant physicians can be evaluated both in general and for each category.

Source of Finance

During this study, no financial or spiritual support was received neither from any pharmaceutical company that has a direct connection with the research subject, nor from a company that provides or produces medical instruments and materials which may negatively affect the evaluation process of this study.

Conflict of Interest

No conflicts of interest between the authors and / or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

All authors contributed equally while this study preparing.

REFERENCES

1. Satterly LB. Teaching differential diagnosis: tool box techniques. J Physician Assist Educ. 2020;31(2):77-84. [Crossref] [PubMed]
2. Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, et al. Competency-based medical education: theory to practice. Med Teach. 2010;32(8):638-45. [Crossref] [PubMed]
3. Zaweski J, Melcher BQ, Sedrak M, Von M, Fletcher S. Physician assistant educator competencies. J Physician Assist Educ. 2019;30(1):47-53. [Crossref] [PubMed]
4. Frye AW, Hemmer PA. Program evaluation models and related theories: AMEE guide no. 67. Med Teach. 2012;34(5):e288-99. [Crossref] [PubMed]
5. Öztürk R. Ülkemizde tıp eğitimine genel bir bakış [An overview of medical education in our country]. Sağlık Düşüncesi ve Tıp Kültürü Platformu Dergisi. 2006-2007;1. [Link]
6. Söylemezoğlu F, Sökmensüer C, Sungur A. Yeterliğe dayalı patoloji uzmanlık eğitimi programı: Hacettepe Üniversitesi deneyimi [Competence-based pathology residency training program: Hacettepe University experience]. Türk Patoloji Dergisi. 2009;25(2):35-40. [Crossref]
7. Levett-Jones T, Gersbach J, Arthur C, Roche J. Implementing a clinical competency assessment model that promotes critical reflection and ensures nursing graduates' readiness for professional practice. Nurse Educ Pract. 2011;11(1):64-9. [Crossref] [PubMed]
8. Epstein RM, Hundert EM. Defining and assessing professional competence. JAMA. 2002;287(2):226-35. [Crossref] [PubMed]
9. Akalın E. Tıpta uzmanlık eğitimi: kalite arayışı [Medical residency training: the pursuit of quality]. Türk Nöroşirürji Dergisi. 2002;12:99-100. [Link]
10. Gültekin BK, Söylemez A, Dereboy İF, Çiçek C. Ege ve Adnan Menderes Tıp Fakültelerinde uzmanlık eğitimi: tıpta uzmanlık öğrencisi bakış açısı ile [Specialization education in Ege and Adnan Menderes Medical Faculties: from the point of view of the medical residency student]. ADÜ Tıp Fakültesi Dergisi. 2006;7(2):17-21. [Link]
11. Anıl D. Tıp eğitiminde ölçme ve değerlendirme [Measurement and evaluation in medical education]. Sağlık Düşüncesi ve Tıp Kültürü Platformu Dergisi. 2014-2015;33:82-7. [Link]
12. Aydın S. Tıpta uzmanlık eğitiminde gelecek öngörüleri [Future predictions in medical residency training]. Sağlık Düşüncesi ve Tıp Kültürü Platformu Dergisi. 2016;40:84-9. [Link]
13. Yalabık HA, Musal B. Tıp eğitiminde klinik dönemde kullanılabilecek değerlendirme yöntemlerinden örnekler [Examples of evaluation methods that can be used in clinical education in medical education]. DEÜ Tıp Fakültesi Dergisi. 2017;31(3):153-67. [Crossref]
14. Gülpınar MA. Klinik eğitimde ölçme-değerlendirme. Aslan D, Özyurt A, editörler. Türk Tabipleri Birliği Uzmanlık Dernekleri Eşgüdüm Kurulu XIV. Tıpta Uzmanlık Eğitimi Kurultayı Bildiri Kitabı. Ankara: Mattek Matbaacılık; 2008. p.29-38.
15. Yılmaz Y, Uçar E, Ertin H. Tıpta uzmanlık eğitimi ve asistan hekimlerin sorunlarının irdelenmesi: bir anket çalışması [Examining the problems of residency training in medicine and residents: a survey study]. Tıp Eğitimi Dünyası. 2019;18(54):21-9. [Crossref]
16. Carr SJ. Assessing clinical competency in medical senior house officers: how and why should we do it? Postgrad Med J. 2004;80(940):63-6. [Crossref] [PubMed] [PMC]
17. Tengiz Fİ, Şahin H. Klînîkte eğîtîmde yenî bîr ölçme yöntemî: mînî klînîk değerlendîrme [A new measurement method in clinical training: the mini-clinical assessment]. Tıp Eğitimi Dünyası. 2014;39:13-8. [Link]
18. Bishop YMM, Feinberg SE, Holland PW. Discrete Multivariate Analysis-Theory and Practice. 1st ed. Cambridge, Mass.: The MIT Press; 1975.
19. Congalton RG, Mead RA. A quantitative method to test for consistency and correctness in photointerpretation. Photogramm Eng Rem S. 1983;49(1):69-74. [Link]
20. Congalton RG, Oderwald RG, Mead RA. Assessing landsat classification accuracy using discrete multivariate analysis statistical techniques. Photogramm Eng Rem S. 1983;49(12):1671-8. [Link]
21. Helldén U. A Test of Landsat-2 Imagery and Digital Data for Thematic Mapping Illustrated by an Environmental Study in Northern Kenya. 1st ed. Report No. 47. Sweden: Laboratory of Remote Sensing, Department of Physical Geography, University of Lund; 1980.
22. Short NM. The Landsat Tutorical Workbook: Basics of Satellite Remote Sensing. 1st ed. Greenbelt, Md,: NASA Reference Publication; 1982.
23. Rosenfield GH, Fitzpatrick-Lins K. A coefficient of agreement as a measure of thematic classification accuracy. Photogramm Eng Rem S. 1986;52(2):223-7. [Link]