Contents lists available at ScienceDirect

Chinese Journal of Traumatology

Original Article

Reliability of trauma coding with ICD-10

Farkhondeh Asadi[*], Maryam Ahmadi Hosseini, Sohrab Almasi

Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

A B S T R A C T

Purpose: The reliability of trauma coding is essential in establishing the reliable trauma data and adopting efficient control and monitoring policies. The present study aimed to determine the reliability of trauma coding in educational hospitals affiliated to Shahid Beheshti University of Medical Sciences, Iran.
Methods: In this descriptive cross-sectional study, 591 coded medical records with a trauma diagnosis in 2018 were selected and recoded by two coders. The reliability of trauma coding was calculated using Cohen's kappa. The data were recorded in a checklist, in which the validity of the content had been confirmed by experts.
Results: The reliability of the coding related to the nature of trauma in research units was 0.75−0.77, indicating moderate reliability. Also, the reliability of the coding of external causes of trauma was 0.57 −0.58, suggesting poor reliability.
Conclusion: The reliability of trauma coding both in terms of the nature of trauma and the external causes of trauma does not have a good status in the research units. This can be due to the complex coding of trauma, poor documentation of the cases, and not studying the entire case. Therefore, holding training courses for coders, offering training on the accurate documentation to other service providers, and periodically auditing the medical coding are recommended.

Introduction

Trauma is a pervasive phenomenon and has become a major global health concern, with negative effects on the individual, family and society.[1,2] The global burden of trauma is very high, with injuries constituting 10% of all the deaths and >15% of disability-adjusted life years.[3,4] Millions of deaths with trauma represent only a small fraction of the tens of millions of injured people, who are hospitalized or visit a general practitioner.[5] Trauma and injuries occur frequently, which are associated with significant mortality. The worst consequence of physical injuries is loss of life. Non-fatal injuries impose heavy personal and social costs.[6] To effectively manage and prevent trauma, it is important to identify the factors that may cause and affect it.[7] The first step in developing countries, a harm-prevention strategy is injury monitoring. To this end, the availability of high-quality data on injury, assessment of the current risk of injury, identification of potential risk factors and establishment of a preventive intervention of injury are crucial.[8] Therefore, accurate information is required on the mechanism of harm to inform prevention programs,[9] and information retrieval standards should be established.

Coding of clinical data is integral to the compatibility and retrieval of standard medical information, and clinical reports depend primarily on coded data. Maintaining the desired quality of coding is significant for the accurate analysis and interpretation of clinically important data.[10] According to Avila-Weil and Regan,[10] the reliability, accuracy and completeness of the code are the most important components of coding quality. Coding reliability refers to obtaining the same results upon repeating the coding activity.[11] In other words, the coding reliability is an agreement between different people in coding a diagnosis (external reliability), or an individual coding the same diagnosis at different time (internal reliability).[12] Adopting a consistent encoding process that results in the reliable coded data is crucial to using these data because users will trust the data when they are convinced that the data encoding process is reliable.[13] The reliability of the data indicates the stability of the data since unreliable information reduces the usefulness of the system.[14]

* Corresponding author.
    E-mail address: asadifar@sbmu.ac.ir (F. Asadi).
    Peer review under responsibility of Chinese Medical Association.

Various studies on the reliability of medical coding have shown different levels of reliability. For instance, Peng et al.[15] reported the International Statistical Classification of Diseases and Related Health Problems − Tenth Revision (ICD-10) coding agreement and reliability of ICD diagnosis codes in emergency department records was 86.5% and 82.2% at 3 and 4 digits levels, respectively, and reliability was 0.86 and 0.82, respectively, between auditors and hospital coders. The main causes of coding discrepancies between coders were the use of codes at different levels of specificity and the utilization of unspecified coding in practice. Daniels[16] also examined the reliability of the coding related to a children's hospital and stated that the reliability of the data is poor. He cited limitations in the ICD, in which ambiguous medical records documentation and inadequate coder training are the reasons for the lack of code reliability.[16]

In a trauma registry system, the data can be used to assess injury management, injury protocols and hospital statistics, only if they are reliable. An unreliable registry can mislead the hospital statistics at both regional and national levels.[17] Considering the importance of trauma prevention and the accurate and reliable data obtained from high-quality trauma code and its external causes, the present study investigated the reliability of trauma coding as a component of trauma coding quality with ICD-10.

Methods

This descriptive cross-sectional study was conducted in 2018 to evaluate the reliability of the coding of trauma patients' records with ICD-10. The research population comprised all coded medical records with the final diagnosis of trauma in 2018 in 8 educational hospitals affiliated to Shahid Beheshti University of Medical Sciences (Iran). To collect the data, a researcher-made checklist was employed to evaluate the reliability of the coding related to the nature of trauma (main condition) and the coding of external causes of trauma (external causes). The validity of the checklist content had been confirmed by experts. To determine the sample size, each educational and medical center was considered as a class.

Using stratified sampling methods and calculating the sample size through Cochran's formula in infinite population and considering $d = 0.05$ and $z = 95\%$ confidence interval, 591 coded medical records were selected as the sample size.

$$n = \frac{z^2 pq}{d^2}$$

In the equation, "n" denotes the sample size, "z" represents the value of the normal variable (1.96), "p" is the disagreement rate, "q" is 1-p and "d" is the significant level (0.05).

All components of medical diagnoses with the coding assigned by the coder working in the research units were recorded in the checklist. The two coders then recoded the diagnosis of the nature of trauma and the external causes of trauma at different times and places. Both coders were health information technology experts with >10 years of experience in the coding unit. After collecting the data via the checklist, they were inputted to SPSS version 20 and analyzed by Cohen's kappa test, which is often utilized to test the reliability. The importance of reliability is that it indicates the degree to which the data collected in the study accurately represents the measured variable. Cohen's kappa is a strong and standard statistical value and the correlation coefficient can vary from -1 to 1, where 1 indicates complete agreement between individuals. Cohen's kappa value < 0 is impossible in practice. The rating provided by McHugh was adopted to judge the status of the reliability coefficient using Cohen's kappa (Table 1).[18]

Table 1
Interpretation of Cohen's kappa values based on McHugh's rating.

Value of kappa	Level of agreement	Percentage of reliable data
0−0.20	None	0%−4%
0.21−0.39	Minimal	4%−15%
0.40−0.59	Weak	15%−35%
0.60−0.79	Moderate	35%−63%
0.80−0.90	Strong	64%−81%
Above 0.90	Almost perfect	82%−100%

Results

The calculation of Cohen's kappa coefficient to examine the reliability of the coding related to the nature of trauma is presented in Figs. 1 and 2. The agreement between Coder 1 and the primary coder on the coding of the nature of trauma in the research units was 0.75, suggesting average reliability (Table 1). The agreement between Coder 2 and the primary coder on the coding of the nature of trauma in the research units was 0.77, which indicated average reliability (Table 1).

In the next step, the agreement between Coders 1 and 2 and the primary coder about the external causes of trauma coding was examined (Figs. 3 and 4). The agreement of Coder 1 with the primary coder on the coding of external causes of trauma in the research units was 0.57 which suggested poor reliability (Table 1). The agreement of Coder 2 with the primary coder on the coding of external causes of trauma in the research units was 0.58 which indicates poor reliability (Table 1).

Discussion

Today, medical coding and its quality are becoming increasingly important due to their relationship with the quality of medical care.[19] The reliability is one of the dimensions of coding quality. The usefulness of classified and coded medical data essentially depends on the uniform coding of similar entities, independent of the coder or coding time.[20] Coding trauma cases is difficult due to the existence of multiple and complex coding rules. In this regard, Curtis et al.[21] stated that accurate coding and evaluation of clinical information coding in trauma patients are integral to trauma management. Coding for trauma patients is generally difficult because providing and documenting medical care are complex processes. Care programs for such patients may comprise multiple types of care during unpredictable hospitalization, which leads to problems in coordinating different care and treatment services. The low quality of trauma patient codes is probably due to the complex nature of trauma, extensive patient records, imperfect documentation, multiple traumas and the use of multiple codes.

The results of the present study revealed that the reliability of the nature of trauma coding in the research units varies from 0.75 to 0.77, indicating the average reliability. As for the external causes of trauma coding in the research units, the findings showed that the reliability is 0.57−0.58, suggesting the poor reliability (Table 1). After investigating the reliability of cause of death coding, Antini et al.[22] reported the reliability of 76.4% for all death and 80.6% for the main coding of death. Misset et al.[23] also evaluated the reliability of coding related to an intensive care unit and stated that the agreement was 34% between two coders and only 18% among three coders. Furthermore, Daniels[16] studied the reliability of the coding related to a children's hospital and concluded that the reliability of coding in this hospital is poor.

Little research has been conducted on the reliability of coding the external causes of trauma. Scott et al.[24] reported the reliability of 0.94 and 0.97 for individual activity at the time of the trauma and
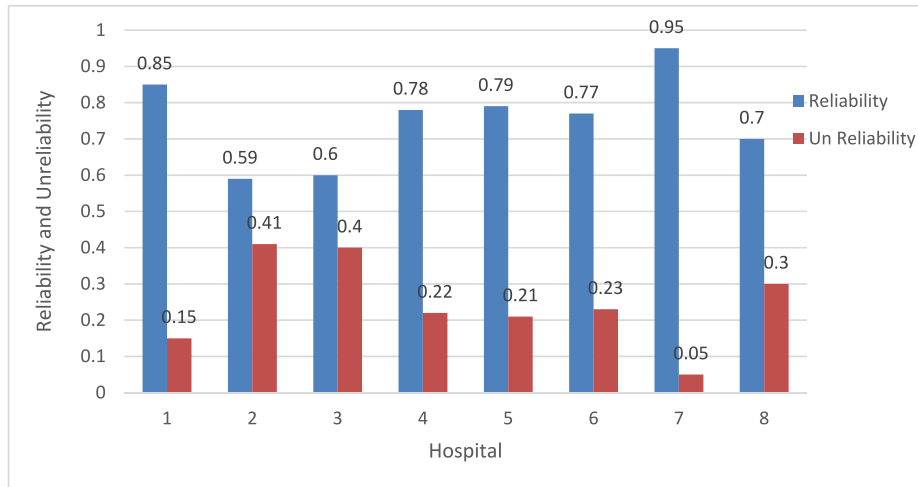
Fig. 1. Agreement between Coder 1 and the primary coder on the reliability of the nature of trauma coding in the research units.
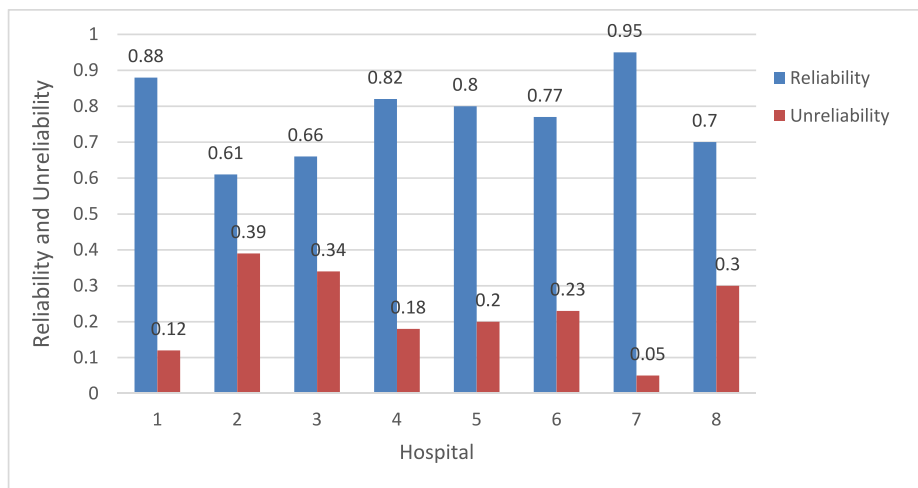

Fig. 2. Agreement between Coder 2 and the primary coder on the reliability of the nature of trauma coding in the research units.
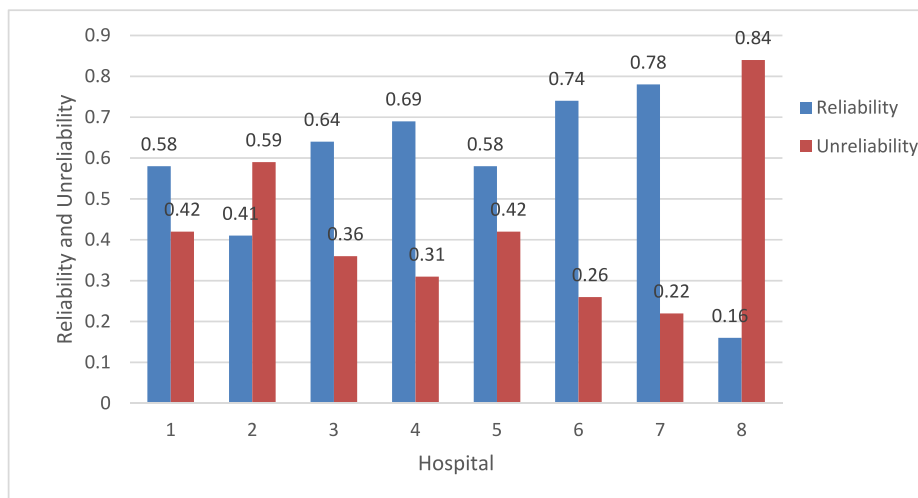

Fig. 3. Agreement between Coder 1 and the primary coder on the external causes of trauma coding in the research units.
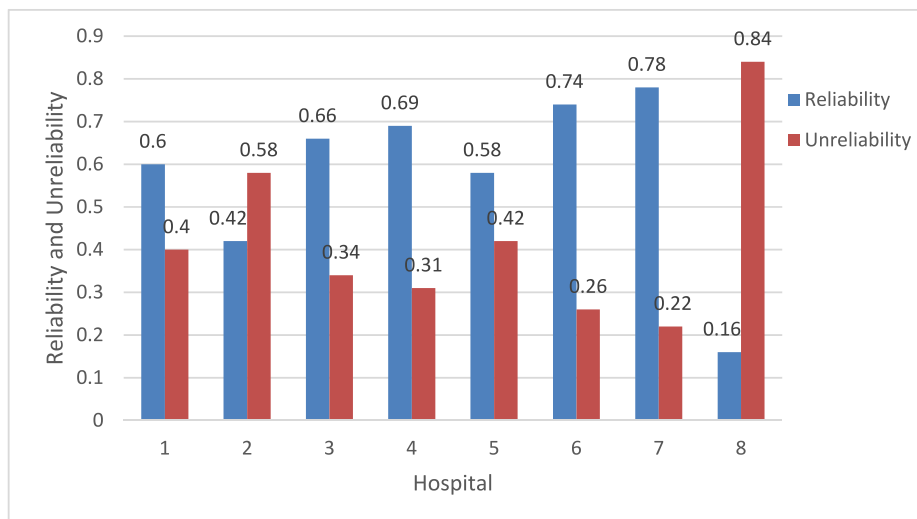
Fig. 4. Agreement between Coder 2 and the primary coder on the coding of external causes of trauma in the research units.

the location of the trauma, respectively. Neil et al.[25] stated that the reliability of the external cause of trauma coding was 0.80, intention (intentional or unintentional trauma) was 0.58, and location of the trauma was 0.44.

In the research units, no accurate and high-quality medical records, no cohesion and coherence in the documents and exited contradictions in the documentation caused the most errors and confusion of the coder. For example, radius trunk fractures were reported in the medical history sheet, but distal radius fractures were recorded in others sheet; the diagnoses were not recorded in the admission and discharge summery sheet. Moreover, the coders coded only by studying the medical history sheet, whereas other sheets and the operation report sheet were not examined. The number of multiple coding, strict rules of coding trauma, and lack of training of coders were the other factors affecting the reliability of coding. Coding rules are forgotten due to the absence of training or ineffective training, which affects the quality of the code. Failure to use the cover of the ICD-10 (Tabular list), read the entire medical records (by the coder), and complete the diagnosis in the admission and discharge summery sheet (by the physician) are other factors leading to the poor reliability of trauma coding. Also, the location of trauma, person's activity during trauma, and whether the trauma is intentional or unintentional has not been recorded in most cases.

According to Annest et al.,[9] the accurate medical records require to contain sufficient detail about the condition of the injury. However, evidence suggests that healthcare providers may not record these details because they see no reason to do so. In addition, sometimes they are unable to determine the details of the situation. For instance, an unconscious person may not be able to provide any information about how the injury has occurred. Stausberg et al.[12] argue that some well-established problems, such as ambiguities and inconsistencies in the ICD, as well as the complexity of coding rules, raise concerns of the current reliability of coding diagnoses with ICD. In a study of Peng et al.,[15] on the reliability of emergency coding, it was stated that the most common difference between coders was that they assigned different codes to the same condition (23.6%). The issue of coding properties is another common cause of discrepancy because one encoder may be assigned to more specific codes.[15] Stausberg et al.[12] believed that the low reliability may have two causes: insufficient training, inadequate standardization of the coder and the coding scenario; weaknesses in the classification system adopted for coding. In the

latter case, the low reliability indicates the poor quality of a coding system and needs fundamental revision. A research also shows that differences in the interpretation of coding rules lead to differences in coding reliability.[22] Using ICD-based administrative data without knowing its regulations may result in misinterpretation. The findings thus highlight the need for standardization of documents.[26]

The issue of coding specificity is another major reason for differences in coding, which can lead to ambiguities in the medical records reported by the coder. Providing an accurate definition of coding, as well as continuous training of coders, are essential for ensuring optimal coding reliability.[15] Thus, the lack of adequate and specialized training for employees should be considered as an important factor.

Studies demonstrate a clear link between poor medical documentation and poor coding reliability. The reasons for the unreliability of coding include: lack of experienced and specialized coding staff, insufficient training of the current coding staff (medical and non-medical), poor medical documentation, lack of training and/or poor usability of the office data software, inherent limitations of the ICD-10 coding system in relation to concepts/terms, the insignificance of coding, and lack of complete regular auditing of administrative data. Therefore, the employment of trained coders at the specialized level and presenting courses on data quality are beneficial interventions for guaranteeing high coding reliability.[16] Studies on coding reliability in other countries report different results. In some of them, a degree of coding reliability has been set, while in some others, the results have been expressed moderate reliability.[12,15,16] These results are in line with the results of the present study.

The findings of the present study showed that the reliability of trauma coding both in terms of the nature of trauma and the external causes of trauma does not have a good status in the research units. The main reasons for these results are the use of general and unclear codes, not reading the entire medical records, and not using the cover of the ICD-10 (Tabular list). The unreliable code can limit the health research and planning projects as it will complicate the statistical analysis of information. Accordingly, periodical evaluating and auditing the codes, training the coders, making the quantitative and qualitative analysis of medical records, and keeping accurate and complete documentation of medical records are recommended.

Funding

Ethical statement

Not applicable.

Declaration of competing interest

The authors have no conflicts of interest relevant to this article to disclose.

Acknowledgements

Author contributions

Farkhondeh Asadi: conceptualization and design, draft manuscript preparation. Maryam Ahmadi Hosseini and Sohrab Almasi: methodology. Farkhondeh Asadi and Maryam Ahmadi Hosseini; data collection. Maryam Ahmadi Hosseini: data analysis and interpretation of results. All authors reviewed the results and approved the final version of the manuscript.

References

1. Magruder KM, Kassam-Adams N, Thoresen S, et al. Prevention and public health approaches to trauma and traumatic stress: a rationale and a call to action. *Eur J Psychotraumatol.* 2016;7:29715. https://doi.org/10.3402/ejpt.v7.29715.
2. Magruder KM, McLaughlin KA, Borbon DLE. Trauma is a public health issue. *Eur J Psychotraumatol.* 2017;8:1375338. https://doi.org/10.1080/20008198.2017.1375338.
3. Asadi F, Paydar S. Presenting an evaluation model of the trauma registry software. *Int J Med Inf.* 2018;112:99–103. https://doi.org/10.1016/j.ijmedinf.2018.01.013.
4. Chichom-Mefire A, Nwanna-Nzewunwa OC, Siysi VV, et al. Key findings from a prospective trauma registry at a regional hospital in Southwest Cameroon. *PloS One.* 2017;12, e0180784. https://doi.org/10.1371/journal.pone.0180784.
5. *Injuries and violence: the facts 2014.* WHO Web Site; 2020. www.who.int/violence injury prevention/key_facts/en/.
6. Duckworth MP, Iezzi T. Physical injuries, pain, and psychological trauma: pathways to disability. *Psychol Inj Law.* 2010;3:241–253. https://doi.org/10.1007/s12207-010-9086-8.
7. Bonilla-Escobar FJ, Gutiérrez MI. Injuries are not accidents: towards a culture of prevention. *Colomb Méd.* 2014;45:132–135.
8. Cross M, Williams S, Kemp SPT, et al. Does the reliability of reporting in injury surveillance studies depend on injury definition? *Orthop J Sports Med.* 2018;6. https://doi.org/10.1177/2325967118760536, 2325967118760536.
9. Annest JL, Fingerhut LA, Gallagher SS, et al. Strategies to improve external cause-of-injury coding in state-based hospital discharge and emergency department data systems; recommendations of the CDC Workgroup for Improvement of External Cause-of-Injury Coding. *MMWR Recomm Rep (Morb Mortal Wkly Rep).* 2008;57:1–15.
10. Nair GJ. Ensuring quality in the coding process: a key differentiator for the accurate interpretation of safety data. *Perspect Clin Res.* 2013;4:181–185. https://doi.org/10.4103/2229-3485.115383.
11. Avila-Weil D, Regan R. *Independent Medical Coding: The Comprehensive Guidebook for Career Success as a Medical Coder.* Rayve Productions; 2006.
12. Stausberg J, Lehmann N, Kaczmarek D, et al. Reliability of diagnoses coding with ICD-10. *Int J Med Inf.* 2008;77:50–57. https://doi.org/10.1016/j.ijmedinf.2006.11.005.
13. Bielby JA. Reliable coded data require a reliable coding process framework. *J AHIMA (Am Health Inf Manag Assoc).* 2014;85:80–82.
14. Greiver M, Barnsley J, Glazier RH, et al. Measuring data reliability for preventive services in electronic medical records. *BMC Health Serv Res.* 2012;12:116. https://doi.org/10.1186/1472-6963-12-116.
15. Peng M, Eastwood C, Boxill A, et al. Coding reliability and agreement of International Classification of disease, 10th revision (ICD-10) codes in emergency department data. *Int J Popul Data Sci.* 2018;3:445. https://doi.org/10.23889/ijpds.v3i1.445.
16. Daniels A. *Examining the Reliability of Discharge ICD Coding in the Red Cross War Memorial Children's Hospital Administrative Database.* University of Cape Town; 2018.
17. Horton E, Krijnen P, Molenaar H, et al. Are the registry data reliable? An audit of a regional trauma registry in The Netherlands. *Int J Qual Health Care.* 2017;29(1):98–103. https://doi.org/10.1093/intqhc/mzw142.
18. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med.* 2012;22(3):276–282. https://doi.org/10.11613/BM.2012.031.
19. Leon-Chisen N. Coding and quality reporting: resolving the discrepancies, finding opportunities. *J AHIMA (Am Health Inf Manag Assoc).* 2007;78(7):26–30. PMID: [17679594].
20. Alipour J, Karimi A, Erfannia L, et al. Reliability of medical diagnosis with international classification of Diseases 10th version in 2011. *Health Inf Manag.* 2013;10(1).
21. Curtis K, Bollard L, Dickson C. Coding errors and the trauma patient-is nursing case management the solution? *Aust Health Rev.* 2002;25(4):73–80. https://doi.org/10.1071/AH020073.
22. Antini C, Rajs D, Muñoz-Quezada MT, et al. Reliability of cause of death coding: an international comparison. *Cad Saúde Pública.* 2015;31:1473–1482. https://doi.org/10.1590/0102-311x00099814.
23. Misset B, Nakache D, Vesin A, et al. Reliability of diagnostic coding in intensive care patients. *Crit Care.* 2008;12:R95. https://doi.org/10.1186/cc6969.
24. Scott D, Harrison J, Purdie D, et al. The properties of the International Classification of the External Cause of Injury when used as an instrument for injury prevention research. *Inj Prev.* 2006;12:253–257. https://doi.org/10.1136/ip.2006.011510.
25. Neale R, Rokkas P, McClure RJ. Interrater reliability of injury coding in the queensland trauma registry. *Emerg Med.* 2003;15:38–41. https://doi.org/10.1046/j.1442-2026.2003.00406.x.
26. Triep K, Beck T, Donzé J, et al. Diagnostic value and reliability of the present-on-admission indicator in different diagnosis groups: pilot study at a Swiss tertiary care center. *BMC Health Serv Res.* 2019;19:23. https://doi.org/10.1186/s12913-018-3858-3.