



Cross-analysis of big data in accreditation of health specialists

Marina B. Chelyshkova¹, Tatyana V. Semenova², Natalia N. Naydenova^{1,3}, Evgeniy M. Dorozhkin⁴, Aleksey A. Malygin⁵, Vasiliy V. Akhunov¹

ABSTRACT

Objective: The relevance of this study is due to the mass accreditation of health professionals that is developing in Russia, which requires innovative measurement tools and opens new opportunities for a well-founded cross-analysis of specialists' professional readiness quality. Purpose of the study: The purpose of this article is to present approved methodical approaches to the transformation of accreditation data into a format suitable for secondary analysis of medical schools graduates quality based on the requirements of Professional Standards.

Method: The leading methods of secondary data analysis are: a) codification of indicators in the primary data accumulation array; b) statistical processing of study results (evaluation of the relationships between the arrays of primary data accumulation and instrumental data, the correlation of test scores obtained by accreditation results with the labor functions of Professional Standards); c) the creation of representative samples for data analysis. The implementation of methods is carried out in the mode of working with arrays of big data, which also uses the method of cross-analysis to identify additional factors that affect to specialists' professional readiness quality.

Results: As a results of the research, there were: 1) approaches to the codification of data in the array and their secondary analysis were developed; 2) three samples were constructed with an estimation of representativeness for different strata, including subjects, assignments and corresponding labor functions; 3) the matrix of primary data in the specialty "Pediatrics" was verified using the example of the results of students from 50 medical universities in Russia.

Conclusion: Approbation of methods of secondary data analysis conducted on representative samples of the subjects showed the effectiveness of the developed approaches that should be used when analyzing large data sets in the procedures of certification or accreditation. The materials of the article can be useful for specialists in the field of assessing the quality of education or assessing the professional readiness of health professionals, managers, professors and pedagogical staff of medical schools, specialists of centers for independent assessment of qualifications.

Keywords: large data, primary and secondary analysis, representative samples, cross-analysis, labor functions

INTRODUCTION

Relevance of the Problem

The primary analysis of the accreditation data, aimed at comparing the scores of the subjects with passing scores at its individual stages, allows only to make decisions on accreditation, but it does not provide any detailed information on the quality of the professional readiness of health professionals. Secondary cross-sectional analysis opens up new opportunities in interpreting the results of accreditation for improving the quality of medical and pharmaceutical education and increasing the professional readiness of health professionals.

At the federal level, the data from secondary analysis will allow the Ministry of Health to make a corrective impact on its decisions based on feedback, based on the structuring and consolidation of information about the quality of medical and pharmaceutical education in Russia. Based on this information, strategic directions for the development of medical education can be outlined, ways of implementing the main directions of educational policy in accordance with the needs of society and the state can be determined. The results of primary and primary specialized accreditation provide

¹ First Moscow State Medical University named after I. M. Sechenov, Moscow, Russia

² Ministry of Health, Moscow, Russia

³ Federal State Budget Scientific Institution "Institute for Strategy of Education", Moscow, Russia

⁴ Russian State Vocational Pedagogical University, Ekaterinburg, Russia

⁵ Ivanovo State University, Ivanovo, Russia

Correspondence: Marina B. Chelyshkova

First Moscow State Medical University named after I. M. Sechenov, Moscow, Russia.

E-mail: mchelyshkova@mail.ru

Received: 21 Apr 2018, Accepted: 8 Jul 2018

© 2018 by the authors; licensee Modestum Ltd., UK. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>).

Electronic Journal of General Medicine

educational information that is an indicator of the state of the educational process in different universities, shows the success of educational programs and the degree to which graduates are fit for the requirements of GEF and Professional Standards.

At the level of individual structures, the secondary analysis of accreditation data provides crucial information both for individualizing the training of specialists in the system of continuous medical education and for improving the effectiveness of educational activities of various educational organizations. Such information will allow the heads of the medical education system to make informed management decisions on replicating the positive experience of the country's leading universities in the health care system, build strategies for helping lagging universities, identify points of growth and establish the reasons for the failure in the training of doctors and pharmacists. To a large extent, the solution of these problems is facilitated by the development of computer support for quality management systems of education, characteristic of the second decade of the 21st century. New and more advanced versions of these systems, based on valid data from mass evaluation processes, enable the analysis of process quality and learning outcomes. In general, the secondary analysis of accreditation results should have a complex cross-type character and be conducted in several areas that are in continuous interaction to establish cause-effect relationships between specialists professional readiness quality and the most significant factors of influence.

The first direction of secondary analysis is intended to improve the measuring instruments themselves in order to increase the reliability and validity of testing data received during tests application in subsequent accreditation procedures. The second area of analysis provides the possibility of using the accreditation results for making managerial decisions in the education system. The third direction is connected with forecasting the success of further specialists work, on which qualifying decisions are made. The fourth direction of accreditation results secondary analysis allows to reveal those requirements of Professional Standards, which are systematically turned out to be poorly developed applicants. The fifth direction of data analysis is comparative research in the system of continuous medical education. The sixth direction is the improvement of Professional Standards requirements and requirements of GEF for vocational education system.

Thus, the results of a comprehensive cross-sectional analysis of accreditation data, which is of secondary nature, are highly relevant and aimed at various consumers, including health authorities, a system of continuing medical education, educational institutions and future health care professionals. However, such complex analysis is complicated not only by the mutual influence of the data consumption spheres, but also by the size of the accreditation data sets. The latter circumstance forces us, in the course of the analysis, to resort to additional data codification, the construction of sample populations and the use of the Big Data regime, which requires innovative methodological approaches to data representation for statistical processing (1). These methodological approaches to the presentation of accreditation data for cross-analysis are described in this article and have several features.

Goals and Objectives of the Study

The purpose of this article is to present methodical approaches to secondary analysis of accreditation data that facilitate their detailed interpretation for improving the professional readiness of health professionals. As the main task, the development of techniques for codifying data, compressing information, constructing representative samples and correlating analysis of data links of the studied population and sample data was chosen. To confirm the effectiveness of the methodologies, their approbation was carried out and the results of the approbation were interpreted.

Status of the Problem Elaboration

The research tasks outlined above are poorly represented in scientific publications and methodological literature on measurement problems, both in general terms and in the field of improving the professional readiness of health professionals. In particular, the publications mainly deal with fundamental or narrowly specialized issues of interpreting information gathered from the results of mass evaluation procedures, but practically do not touch upon methodical approaches to information compression, constructing representative samples and estimating the correlation for cross-analysis of data in Big Data format. Most studies of the results of evaluation procedures are limited to the primary analysis, when the percentages of execution and the average scores for the groups of subjects are calculated on the data matrix without taking into account the factors of influence and the possibilities of additional secondary analysis of the data sets.

Theoretical and Practical Contribution of the Materials of the Article

The results of this research, both theoretical and practical, can include methodological approaches to data codification, data compression, its aggregation and interpretation. Approbation of approaches was carried out on the results of accreditation of graduates of 50 medical universities in Russia in 2017, integrated in federal districts.

According to the approbation data, a correlation analysis was constructed that establishes a measure of the correlation of the investigated indicators, estimated for selective populations and for the general population. Interpretation of the results of approbation was carried out in connection with professional functions in the profession "Pediatrics".

Analysis of Russian Scientific and Pedagogical Literature

Despite the relatively short period of existence of accreditation of health care professionals in Russia, a number of publications have been devoted to it. A number of authors write about the problems of accreditation of health professionals (2). The authors consider a wide range of theoretical and methodological issues, covering the design of instruments, scaling and processing data in order to ensure the high quality of measurement results (3).

The problems of the quality of continuous medical education and its evaluation are considered in the articles of other authors (4); however, all the approaches proposed in the articles to the quality assessment lie, as a rule, outside the methods of the theory of pedagogical measurements, and therefore no one do not address the issues of secondary analysis of data on the quality of medical education.

Most of the articles of other Russian authors are devoted to procedural issues of accreditation (5). However, there are practically no publications in education aimed at secondary analysis of large data sets that provide a detailed information for improving the quality of educational outcomes.

Partially, the problems of secondary analysis of large data sets are investigated by specialists in other fields of science. For example, a number of authors dealt with these problems in sociological studies. (6-8). Basically these publications touch upon the questions of the form of presenting the data of analysis, data collection and storage, but practically do not cover the methods of constructing representative samples of subjects in accordance with modern requirements of measurement theory. Thus, the issues considered in this article are practically not represented in domestic publications on problems of accreditation of specialists or attestation of students.

Analysis of Foreign Studies

The analysis of foreign authors publications of a number of countries that have developed testing systems for certification, and accreditation testifies to the predominant fundamental nature of the research. As a rule, in numerous textbooks and monographs, the basic issues of measuring instruments development and application in assessment and accreditation based on classical or modern test theory are discussed, methodical approaches are proposed to the evaluation of measurement results quality and their scaling (9-12)

The problem of primary and secondary interpretation of data is mainly addressed by the authors of articles, considering it in the context of analyzing the results of certification, certification or accreditation in a certain field of professional activity (13-17). In particular, a number of authors of the articles discuss in a fragmented manner the interpretation and improvement of the quality of evaluation process results in relation to the problem of improving the professional performance of health care professionals (18-25).

In general, it can be concluded that in addition to the fundamental problems of measurement theory, a number of related studies are adjacent to the problems of secondary data analysis, but the issues of secondary analysis connected with the influence of various factors on the results of measurement in assessing the professional readiness of health professionals are insufficiently studied character.

MATERIALS AND METHODS

Theoretical and Empirical Methods

To test the hypothesis about the effectiveness of the methodological approaches proposed by the results of the study, a set of various methods complementary to each other was used:

- theoretical methods based on the analysis of the work of teachers and testers on the research problem, substantiating the methods of working with data in text form on Big Data sets;

- empirical methods, including the codification of indicators in the array of primary data accumulation, statistical processing of research results.

To identify additional factors affecting the quality of professional preparedness of health care professionals, cross-analysis was used in the mode of working with Big Data sets.

Base of Research

The study was conducted on the basis of the Methodological Center for Accreditation of Specialists of the First Medical University named after. Sechenov.

Stages of Research

The study was conducted in three stages. At the first stage, variables (constructs) were allocated for analysis based on primary data sets of accreditation data. At the second stage, the theoretical base of the research was created and corrected. At the third stage, the analysis was carried out, the data were summarized and systematized after identification and aggregation, tabular and graphical representations of the data were calculated. Also at the third stage, samples and additional arrays for secondary analysis were formed on the basis of estimating the correlation of labor functions and test results.

Evaluation Criteria

As qualitative criteria in the empirical part of the study, labor functions of professional standards in the specialty "Pediatrics" were used. The passing quantitative criterion in the testing was chosen at the level of 70 %. Thus, all subjects who fulfilled at the first stage of accreditation at least 70% were considered to have successfully passed the first stage.

Progress and Description of the Experiment

For the experimental part of the study, the results of more than 5,000 graduates of the specialty "Pediatrics" were used, accredited in 2017 in medical schools in Russia. As the data for the study, the results of the first stage of accreditation, including graduate scores for standardized testing, were selected.

The experimental part of the study included several stages. At the first stage, additional variables were introduced for the primary data sets. Then the information was compressed and the introduction of identifiers for the subjects, universities and districts. The next stage of the experiment was devoted to the analysis of the representativeness of the data to confirm the validity of the conclusions and the possibility of their generalization by spreading to the general sets of subjects and tasks. Further, tabular and graphical representations of the data were calculated and additional arrays were created for secondary analysis of the correlation of labor functions and test results using the SPSS package for statistical analysis of data in the social sciences. Then, the results of the processing were interpreted to obtain the conclusions of the study.

RESULTS

Coding and Compression of Information

To codify the matrix of primary data, a new variable was introduced, which was chosen as the federal district code, which allowed to consolidate the statistics and generate a sample of data for the federal districts. Also, all educational organizations were coded with a four-digit code containing some identification information. The identifier of the subject was represented by the electronic address indicated by him at the time of testing.

To compress information from 3112 tasks involved in the accreditation for testing, a representative sample of tasks for cross-analysis was compiled. The sample included 60% of tasks, each of which was not duplicated in files and belonged to one labor function. Further, it was proved that the sample of tasks for cross-analysis with a probability of 99% significantly correlates with the entire data matrix with a coefficient of almost 0.9. Thus, it was concluded that for cross-analysis it is possible to use a sample of tasks and distribute sample data to the whole general set of tasks without loss of reliability and reliability of the results of the analysis.

Formation of Representative Samples

To minimize the volume of data, they were grouped in 8 federal districts (Central Federal District - CFD, North-West Federal District - NWFD, North Caucasus Federal District - NCFD, South Federal District - SFD, Volga Federal District -

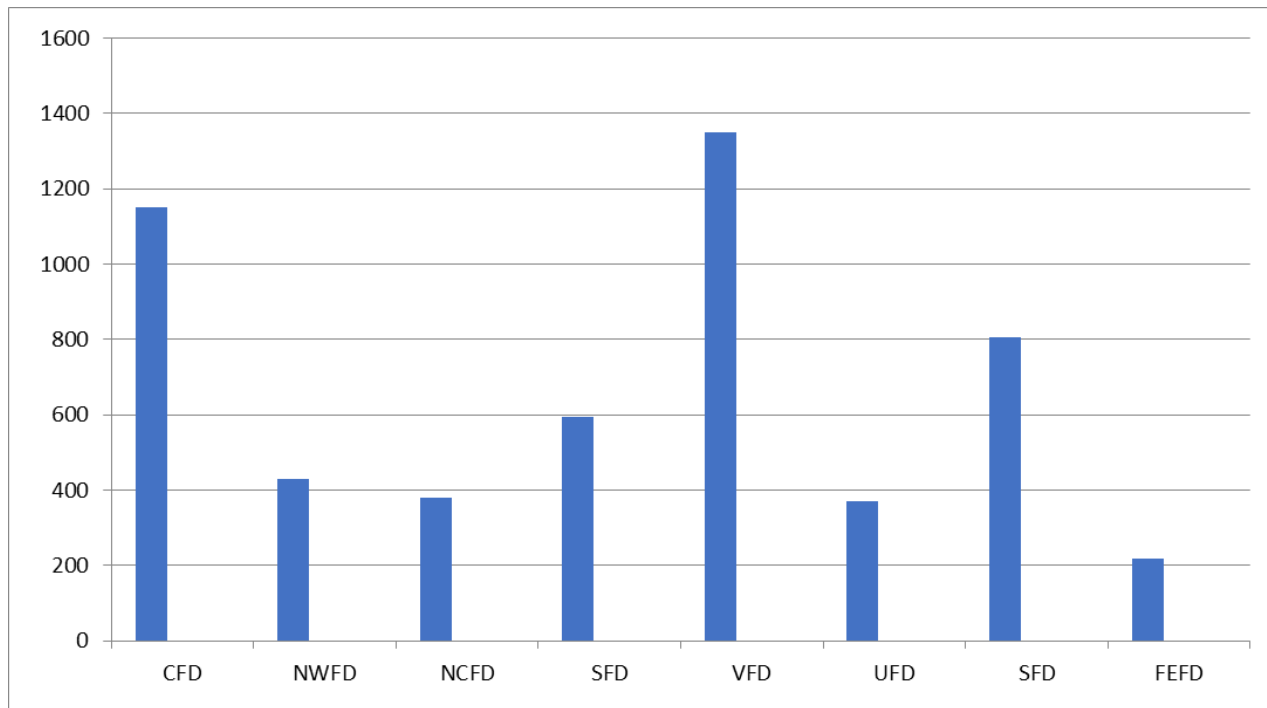


Figure 1: Number of graduates involved in the study, by federal districts

Table 1: Distribution of tasks in the sample according to labor functions (specialty "Pediatrics")

| The code of the labor function in the experiment | The code of the labor function in standard | Purpose of labor function | Number of items | Percent |
|--|--|---------------------------|-----------------|--------------|
| 1 | A/01.7 | Examination | 1598 | 51,3 |
| 2 | A/02.7 | Purpose of treatment | 595 | 19,1 |
| 3 | A/03.7 | Rehabilitation | 121 | 3,9 |
| 4 | A/04.7 | Prevention | 557 | 17,9 |
| 5 | A/05.7 | Management | 241 | 7,7 |
| Total | | | 3112 | 100,0 |

VFD, Ural Federal District - UFD, Siberian Federal District - SFD, Far East Federal District - FEFD). Due to this, the cross-analysis was carried out on the grouped samples shown by **Figure 1**.

The figure shows samples of subjects from 52 universities from 8 federal districts of the Russian Federation. In order to preserve the confidentiality of the results of accreditation for individual HEIs, only the regions are given in the figure, the data on which allow the creation of a matrix for cross-analysis of labor functions.

Cross-analysis of Results on Labor Functions

The results of a qualitative analysis of the sample of tasks conducted to determine the distribution of assignments for labor functions are presented in **Table 1**.

From **Table 1** it follows that the most significant authors of tests are the first function, which is focused on the content of more than 1500 tasks. The second place in importance is occupied by two functions aimed at the treatment and prevention of diseases. Secondary were the third and fifth functions, which included the least number of tasks in the test cases. Since cross-analysis is conducted on compressed, truncated and truncated sample data, which reduces already fragmentary data, the results of the analysis of the last two functions require additional studies to assess the validity and reliability of the results obtained.

Figure 2 shows the results of comparing the average scores obtained from 8 districts based on the compressed cross-analysis data (bottom points) and the total data set (top points).

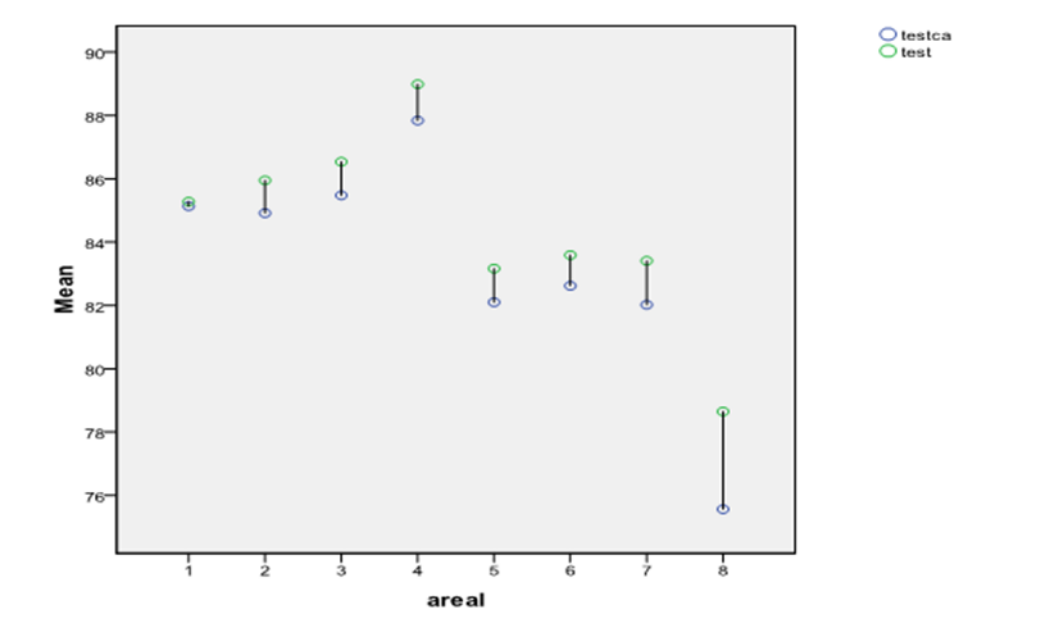


Figure 2: Comparison of average scores for two sets of data

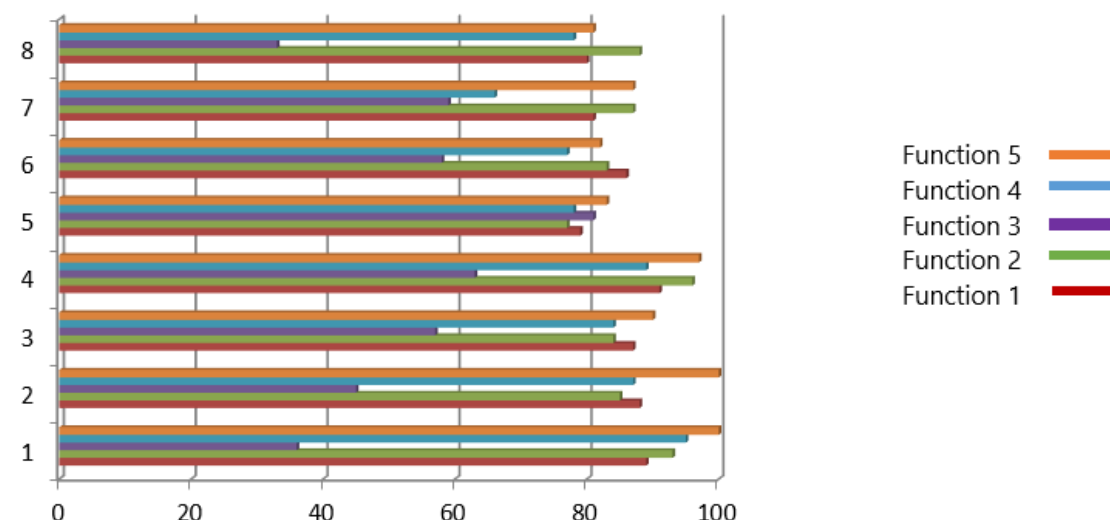


Figure 3: Histogram of average scores for labor functions in districts

The figure introduces the digital notation for the federal districts: 1-CFD, 2 - NWFD, 3 - NCFD, 4 - SFD, 5 - VFD, 6 - UFD, 7 - SFD, 8 – FEFD on the horizontal axis. In the Central District (on the horizontal axis, point 1), there is a complete coincidence of the average scores obtained from the two sets of data. The greatest discrepancy of average scores is observed in the Far Eastern Federal District (on the horizontal axis point 8), but even this seemingly great difference does not exceed two points within the measurement error. Thus, data compression does not lead to loss or distortion of the results of the test subjects and can be used for cross-analysis. In addition to the conclusion drawn in **Figure 2**, another, equally important conclusion follows: the highest average score is observed in the Southern Federal District (on the horizontal axis point 4), and the lowest - in the east of Russia (on the horizontal axis point 8).

The analysis of the Bonferroni analysis of variance showed a 95% statistically significant difference in the average scores of higher education institutions according to the labor functions of professional standards (specialty "Pediatrics") obtained in various federal districts. **Figure 3** shows the histogram of the average scores allocated by color for various labor functions of professional standards in the chosen specialty. The numbers on the left on the vertical axis, as before, correspond to the federal districts.

The analysis of the histogram based on the results of the primary accreditation of graduates opens up opportunities for obtaining a number of conclusions detailing the development of the work functions of graduates and presented in the next section.

DISCUSSIONS

From **Figure 3** it follows that the first function aimed at examining children for the purpose of establishing a diagnosis, in almost all districts, fairly high average scores are observed. A similar even more favorable picture is observed for the fifth function, which requires graduates to organize the activity of medical personnel and maintain documentation. The worst situation is with the development of the third function related to the implementation of rehabilitation programs and the implementation of preventive measures.

Of course, according to the plan of the research, the conclusions obtained are illustrated by the methods of secondary analysis of accreditation data themselves, but cannot at the moment serve as a basis for making managerial decisions in medical education. A detailed and in-depth analysis of additional factors affecting the results of secondary analysis is required. In particular, it is necessary to take into account the number of tasks for each function that are in the sample for cross-analysis, their difficulty and distribution according to the subjects with different levels of professional readiness to perform labor functions. Only if these and a number of other, less important factors are taken into account, the results obtained can serve as the basis for making reliable and valid management decisions.

In general, during the research, the assigned tasks aimed at developing approaches to codification and data compression for their secondary analysis were solved. In the course of solving problems, samples of subjects and assignments with an estimate of representativeness for different strata were constructed and labor functions corresponding to them were identified. Proposed methods of the study were tested on a verified matrix of primary data on the specialty "Pediatrics", which led to the conclusion about the effectiveness of the developed approaches that should be used when analyzing large data sets in accreditation procedures when conducting secondary data analysis.

CONCLUSIONS

As a perspective direction of the development of the considered methods, it is possible to propose an expansion of the list of information on tasks in the process of forming options on the specifications of valuation means. In particular, it is necessary to take into account not only the binding of assignments to certain labor functions, but also their difficulty, determined on representative samples of subjects by statistical means. This will allow obtaining more accurate results on a real data set during secondary analysis and preparation of conclusions for making managerial decisions.

During the cross-analysis of data in the study, approaches to codification and data compression were developed. The formation of samples on two arrays (subjects and tasks) was carried out by the probabilistic-proportional method, which makes it possible to resort to modern methods of analyzing large arrays of accreditation data. The developed approaches to data compression for cross-analysis were supplemented by the dispersion analysis to assess the significance of the differences in the average scores for the federal districts.

The materials of this article can be useful for managers at both the federal and regional levels in order to develop sound decisions in the field of training medical personnel. The developed methods can be useful for teachers of higher educational institutions, testers, analysts and developers of evaluation tools for accreditation or certification of specialists.

REFERENCES

1. Pollak RA. Standardized measurement. Paper prepared for the Workshop on Advancing Social Science Theory: The Importance of Common Metrics. National Academies, Washington; 2010.
2. Semenova T, Sizova Zh, Chelyshkova M, Dorozhkin E, Malygin A. Fairness and Quality of Data in Healthcare Professionals' Accreditation. *Modern Journal of Language Teaching Methods (MJLTM)*. 2017;7:13–25
3. Semenova T, Sizova Zh, Zvonnikov V, Masalimova A, Ersozlu Z. The Development of Model and Measuring Tool for Specialists Accreditation. *EURASIA J. Math., Sci Tech. Ed.* 2017;13(10):6779–788.
4. Dorozhkin EM, Chelyshkova MB, Malygin AA, Toymentseva IA, Anopchenko TY. Innovative Approaches to Increasing the Student Assessment Procedures Effectiveness. *International Journal of Environmental and Science Education*. 2016;11(14):7129-144.
5. Davydova EB. Accreditation to replace certification. Human resources department public (municipal) institution. 2016;11:325-336.
6. Volkova PA, Shipunov AB. Statistical processing of data in educational and research works. Moscow: Forum; 2012.

7. Kryshchanovskiy AO. Use of archive of sociological researches for realization of the secondary and comparative analysis. *Methodological and methods aspects of comparative sociological research*. 1984;2:149-261.
8. Glantz S. *Medical and Biological Statistics. Practica*. Moscow: Academia; 1998.
9. Cizek GJ. *Setting performance standards: Concepts, methods, and Perspectives*. Mahwah: Lawrence Erlbaum Associates; 2011
10. Archer J, Lynn N, Coombes L, Roberts M, Gale T, Price T, Regan de Bere S. The impact of Large Scale Licensing Examinations in highly developed countries: a systematic review. *BMC Medical Education*. 2016;16:212-227. <https://doi.org/10.1186/s12909-016-0729-7> PMID:27543269 PMCID:PMC4992286
11. Lane S, Raymond MR, Haladyna TM. *Handbook of Test Development*. New York: Taylor & Francis; 2014
12. McKinley RK, Fraser RC, Van Der Vleuten C, Hastings AM. Formative assessment of the consultation performance of medical students in the setting of general practice using a modified version of the Leicester Assessment Package. *Medical Education*. 2000;34:573-579. <https://doi.org/10.1046/j.1365-2923.2000.00490.x> PMID:10886641
13. Gessmann HW, Sheronov EA. Psychological Test Validity. *Journal of Modern Foreign Psychology*. 2013;2(4):20-31.
14. Harrison CJ, Könings KD, Schuwirth LW, Wass V, van der Vleuten C. Changing the Culture of Assessment: the dominance of the summative assessment paradigm. *BMC Medical Education BMC series – open, inclusive and trusted*. 2017;17:73-86. <https://doi.org/10.1186/s12909-017-0912-5>
15. De Ayala R. *The Theory and Practice of Item Response Theory*. London: Guilford Press; 2009.
16. Koike S, Masatoshi Matsumoto, Hiroo Ide, Hideaki Kawaguchi, Masahisa Shimpo and Hideo Yasunaga. Internal medicine board certification and career pathways in Japan. *BMC Medical Education. BMC series – open, inclusive and trusted*. 2017;17:83-97.
17. Jonson R, Penny J, Gordon B. *Assessing Performance. Designing, Scoring, and Validating Performance Tasks*. London: Guilford Press; 2009.
18. Baig LA, Violato C. Temporal stability of objective structured clinical exams: a longitudinal study employing item response theory. *BMC Medical Education*. 2012;12:121-132. <https://doi.org/10.1186/1472-6920-12-121> PMID:23216816 PMCID:PMC3552978
19. Carraccio C, Hicks P. *Assessment in Graduate Medical Education: A Primer for Pediatric Program Directors*. Chapel Hill: American Board of Pediatrics; 2011.
20. Edelbring S. Measuring strategies for learning regulation in medical education: Scale reliability and dimensionality in a Swedish sample. *BMC Medical Education*. 2012;12:76-87. <https://doi.org/10.1186/1472-6920-12-76> PMID:22894604 PMCID:PMC3502389
21. Gorter R, Fox J, Twisk J. Why item response theory should be used for longitudinal questionnaire data analysis in medical research. *BMC Med Res Methodol*. Published online; 2015.
22. Govaerts MJB, Van der Vleuten CP, Schuwirth LW, Muijtjens AM. Broadening Perspectives on Clinical Performance Assessment: Rethinking the Nature of In-training Assessment. *Advances in Health Sciences Education*. 2007;12:239-360. <https://doi.org/10.1007/s10459-006-9043-1> PMID:17096207
23. Heeneman S, Oudkerk PA, Schuwirth LT. Department of Pathology, Maastricht The impact of programmatic assessment on student learning: theory versus practice. *Medical Education*. 2015;49:487-498. <https://doi.org/10.1111/medu.12645> PMID:25924124
24. Kersting M, Hornke LF. Improving the quality of proficiency assessment: the German standardization approach. *Psychology Science*. 2006;48(1):85-98.
25. Shulruf B, Wilkinson T, Weller J, Jones P, Poole P. Insights into the Angoff Method: results from a simulation study. *BMC Medical Education. BMC series – open, inclusive and trusted*. 2016;16:134-145.



<http://www.ejgm.co.uk>