

STABILITY OF HOSPITAL RANKINGS

AGATA SIELSKA*

Department of Applied Economics, Warsaw School of Economics,
al. Niepodległości 162, 02-554 Warszawa, Poland

Many hospital rankings are based on algorithms and weights elicited by experts. The paper attempts to build rankings of Polish district hospitals using the TOPSIS method and to examine the sensitivity of the results to the changes in weights. We considered 11 large and 34 medium-sized hospitals. The criteria set consists of man-days total, profit/loss on sales, contract with the National Health Fund, return on assets (ROA) and return on equity (ROE). Because of this, rankings consider different spheres in which hospitals perform, including the financial aspect and their main goal, i.e., treating patients. The results show that despite the overall high similarity of rankings, the benchmarking based on rankings should be done with care as the positions of some individual hospitals changed to a great degree.

Key words: *rankings, stability, hospital, TOPSIS*

1. Introduction

As a result of the unfavourable situation of Polish district hospitals [22], there is a still growing demand for the analysis of their relative status. Multicriteria rankings allow for comparison of a chosen object against the others and, in consequence, they enable effective reaction if the evaluation is unfavourable. The poor relative situation of a hospital may be seen as a need to reallocate the resources [4]. Local authorities are also interested in the performance of the hospital, not only as a means to provide health services for the local population but also as an employer and a way to strengthen the position of a district as a territorial unit. To the best of the author's knowledge, this is the first study in Poland based on such dataset which may, because of the sensitivity analysis, provide an insight in the level of the ranking and its adequacy to the expectations of various decision-makers. Ligsma et al. [18] have already pointed out both the

*Email address: asiels@sgh.waw.pl

Received 3 April 2020, accepted 4 December 2020

popularity of rankings and usefulness of such comparisons for decision-makers in hospitals and potential patients. It seems that rankings should be constructed as based on the criteria which describe both quality and efficiency of hospitals. Often, apart from hard criteria (like efficiency), soft ones are introduced. They can be related among other things to patients' impressions, communication between them and the doctors [30], or communication within the organisation itself [29]. Other factors may affect the ranking, for example, the hospitals' use of social media, as pointed out by Triemstra et al. [33].

Usually, hospital rankings are built on various sets of criteria, the weights of which are determined by experts. For example, the Polish Center for Monitoring the Quality in Health Care¹ has been publishing rankings of hospitals for over 10 years. Rankings are based on criteria related to management, medical care and quality of care [39]. Ranking of hospital wards was published in 2018 by the weekly *Wprost* [40]. The hospital ranking constructed based on financial criteria is presented by Deloitte [41]. Similar comparisons are published in other countries as well, but ranking results can be controversial [1]. This stems from the weights of criteria or some parts of the algorithm. For example, the cited author points out the possible lack of stability of results. Davis et al. refer to potential inconsistencies of hospital assessments in various areas [11]. In general, this phenomenon is to be expected because the criteria, especially those used in the analysis of complicated structures, are often conflicting. Attempts to make holistic assessment are therefore even more important, and multicriteria methods may be used for this purpose.

The main goal of the paper is to build a ranking of Polish district hospitals and to analyse the impact of the change in criteria weights on final results. Rankings are built using multicriteria TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) method and selected weighting algorithms. The paper consists of 3 parts. The first one discusses the features of the TOPSIS method and areas of its application in issues related to the functioning of hospitals. The second part focuses on the methodology, while the third part is dedicated to the presentation and discussion of the results.

2. TOPSIS method in issues related to the functioning of hospitals

TOPSIS [15] is a multicriteria method used for supporting the decision-making process. The construction of ranking begins with the identification of scores of each alternative (hospital) on each criterion. The criteria may be measured on different scales, but the algorithm ensures comparability by normalisation of the assessments

¹Centrum Monitorowania Jakości w Ochronie Zdrowia.

$$r_{ij} = \frac{f_j(a_i)}{\sqrt{\sum_{i=1}^n f_j(a_i)^2}} \quad (1)$$

where $f_j(a_i)$ denotes the score of i th alternative on j th criterion.

Normalised values are multiplied by weights w_j to reflect the relative importance of criteria

$$t_{ij} = f_j(a_i)w_j \quad (2)$$

Alternatives are compared with two reference points. The first one is the hypothetical ideal solution (an alternative that is evaluated best on all criteria), and the second one is the negative ideal, i.e., the hypothetical alternative that is evaluated worst on all criteria. The alternative is assessed as the better, the more distant it is from the negative ideal and the closer to the ideal at the same time. Respective distances are measured in Euclidean metrics. The final ranking is built as based on the relation of the distance from the negative ideal to the sum of distances to both reference points

$$D_p(a_i) = \frac{d_p^-(a_i)}{d_p^-(a_i) + d_p^+(a_i)} \quad (3)$$

where: $d_p^-(a_i)$ – distance of alternative a_i to the negative ideal, $d_p^+(a_i)$ – distance of alternative a_i to the ideal.

TOPSIS is chosen due to its characteristic. For ranking users, it is important to know how the ranking was created and to understand the algorithm or the intuition behind it. TOPSIS fulfils this requirement. It can be assumed that one of the main reasons for the popularity of this method is the ease of use combined with intuitive interpretation, which is based on the concept of distance. The method resembles the approaches used in multivariate comparative analysis, including Hellwig's method [13]. TOPSIS combines the aim to avoid a negative solution with the aim for a positive, while in Simple Additive Weighting method (SAW) the problem is focused only on the weighted average score of alternatives [10]. We did not choose the Analytic Hierarchy Process (AHP) [28], because the weights used in the algorithm are derived from pairwise comparisons of criteria which uses 9-point rating scale (Saaty's scale), and the latter scale does not comply with the approaches used in the stability analysis in this paper. As compared with methods from so-called European School, like ELECTRE (Elimination et Choix traduisant la réalité) [27] and PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluations) [6], TOPSIS is less subjective. ELECTRE and PROMETHEE rely more on

decision-makers preferences (or analyst's preferences in the case when no decision-maker is uniquely defined, as in case of rankings of enterprises). For example, the methods in question use not only weights, but also preference, indifference or veto thresholds (ELECTRE), or generalised criteria functions (PROMETHEE). Because of that, the ranking can be adjusted to the recipient's expectations to a large extent, but as a consequence, the obtained result is highly subjective. Moreover, PROMETHEE allows for incomparability of criteria which, in the author's opinion, does not occur in the case studied in the paper. As pointed out by Nermend [21], TOPSIS features, especially the limited influence of the decision-maker, allow it to be used in multivariate comparative analysis, because the solution is relatively independent of the parameter values adopted during the analysis. From our point of view, the objectivity of the method is one of the most important factors that should be taken into account while building rankings for (supposedly) large audience.

TOPSIS is used in several areas in issues related both to healthcare in general, and functioning of hospitals. The first sphere is related to the assessment of service quality. The quality of hospital services is the subject of several works [8, 20, 32], which is not surprising as this is probably the most important factor from the patients' perspective. It is also important for managers, however, they also need to take into account the economic factors which enable keeping or even improving this quality. For example, Váchová and Hajdíková [35] rank hospitals as based on the economic criteria related to the structure of costs, debt, sales, and productivity of wages. In [30] three hospitals are ranked as based on 6 criteria (with sub-criteria including the willingness to help patients, ability to communicate with them, security). The authors also compare the criteria of importance between different wards. Shafii et al. [29] assess managerial staff as based on, among others, organising, creativity, ability to make adequate decisions, and communication.

Araujo et al. [5] combine TOPSIS with neural networks for the analysis of the healthcare system in Brazil. Their goal was to construct a model on which predictions could be made using macroeconomic or demographical factors. A similar approach is used in [37] to assess the efficiency of diabetes treatment. Du et al. [12] examine TOPSIS combined with DEA, thus taking into account the efficiency factor. Wang and Chou [36] analyse the safety of patients including factors related, i.a., to communication, quality of care, and commitment. Qi et al. [24] evaluate 27 wards, focusing on performance, quality, and economic benefits. Hosseini et al. [14] address the disaster preparedness topic. TOPSIS is also used in problems related to supplier selection [3]. These areas are not related only to hospitals, they are analogous to other industries.

TOPSIS is often combined with AHP which can be applied, for example, to elicit criteria weights [3, 7, 20, 25, 29–30]. Being a well-known method, TOPSIS is also used as a benchmark. For example, Afful-Dadzie et al. [2] in their study of quality health information on the Internet use modification of TOPSIS for comparison with VIKOR, while Tuzkaya et al. with PROMETHEE [34].

To the best of the author's knowledge, the TOPSIS is not used in the analysis of issues related to hospitals in Poland. The reasons may originate from the difficult access to data and the lack of their uniform format. Publishers of hospital rankings use data from surveys. The results are widely available and can affect the patients' opinions and decisions. Hospitals are not always interested in providing data for the needs of scientific work because it involves additional burden for the staff and may not translate into the functioning of the given hospital.

3. Data and methodology

3.1. Data and software

Data used in the study come from the questionnaire prepared by the Polish Association of Employers of Poviats Hospitals² and filled by ca. 110 district hospitals in 2019. The study uses data from the year 2018 which is the last year for which the survey covers all 12 months. Due to the differences between hospitals with different characteristics [22], only public hospitals with emergency ward were included in the study. As based on mean and standard deviation calculated for the number of beds, hospitals are divided into two groups: medium-sized (134–362 beds, i.e., within one standard deviation away from the mean) and large ones (more than 362 beds, i.e., more than one standard deviation away from the mean). Hospitals with a larger number of beds are more likely to have many different wards, better access to financing, but also higher costs. After eliminating cases with missing data, 11 large and 34 medium-sized hospitals remained in the study.

In the remainder of the paper, individual hospitals will be denoted by codenames, i.e., letter L with consecutive numbers will be used for large hospitals, and letter M and consecutive numbers will be used for medium-sized hospitals. Rankings were built using R Software [42] and TOPSIS package [43].

3.2. Criteria

In the paper, hospitals are ranked based on 5 objective criteria. We decided to use variables related to different aspects of hospitals performance. The first criterion, *mandays total*, is related to the effective performance of a hospital and the effects of its core activity. Other criteria are as follows: *profit/loss on sales*, *contract with NHF* (National Health Fund³ – this criterion relates to resources), return on assets (*ROA*) and return on equity (*ROE*).

²Ogólnopolski Związek Pracodawców Szpitali Powiatowych – OZPSP.

³Narodowy Fundusz Zdrowia.

Man-days total and contract with NHF are important because they describe the hospital's ability to fulfil its goals. In expert opinion, profit/loss of sales is one of the most important factors that reflects the overall situation of a hospital.

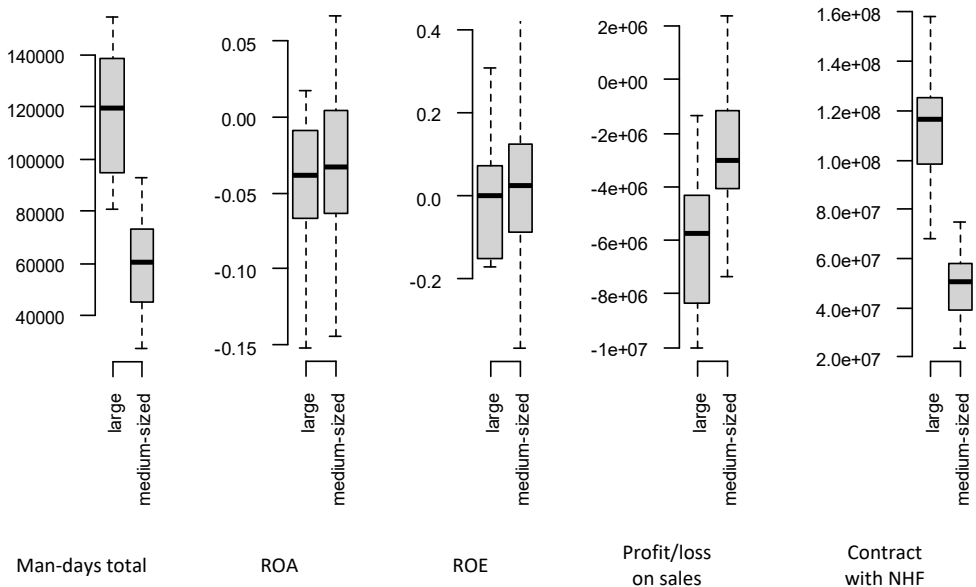


Fig. 1. Distribution of criteria in the analysed groups of hospitals. Based on the data of district hospitals

The pursuit of profitability is controversial in the case of hospitals because of the nature of services they supply but, on the other hand, the profitability of hospitals is in a way forced by their environment [19]. Profitability works in favour of hospitals, allowing them to supply the local population with better services or to make access to treatment easier. Referring to the literature Prędkiewicz et al. [23] point out that the profits open possibilities to improve equipment. They also refer to fact that increasing costs affect the profitability. The reason stressed in the cited paper is related to the patients' and doctors' actions – which cannot be denied – but nowadays the pressure takes the form of the increasing wages. At the moment it is mostly related to the increase of the minimum wage which took place in Poland in January 2020, however, its impact is indirect as well. The main problem of hospitals is not the increase in wages for those who earn a minimum wage but the increase in total costs of employment resulting from internal payroll regulations and rising costs of outsourcing. The importance of profitability indicators is confirmed by the fact that they were included in the draft ordinance of the Minister of Health on monitoring and assessment of the economic and financial situation of hospitals included in the hospital network [9]. Profitability alone is not a sufficient basis for comparison, not only due to the specific goal of hospitals but also due to financing. For example, some commercialised hospitals may focus on selected highly-priced

treatments, while in general public district hospitals should ensure access to healthcare in different dimensions.

Profitability indicators show the hospital's effectiveness in terms of managing equity and assets. Striving to maintain the readability of the paper, in particular in the section focused on sensitivity analysis, only two basic indicators: ROA, and ROE are chosen from the numerous set of profitability indicators which can be used to assess the situation of hospitals. They were used for example in [23, 38].

Ranking results take into account both the financial aspect of the hospital activity and its main goal – treating patients. Due to the character of the data and focus on objective criteria, there was no possibility to include the quality of care. All the criteria are profit (maximised) criteria.

Boxplot comparison of the distribution of criteria in the analysed groups of hospitals is presented in Fig. 1. To improve both the readability of the plots and to increase the anonymity of hospitals, the outliers are not presented.

3.3. Stability of rankings

3.3.1. Scope of stability analysis and weighting approaches

The stability was assessed in two approaches. In the first one, we analyse the changes in the top positions in rankings. Maintaining high rank despite weight changes is important from the point of view of patients, analysts, and hospital managers, because hospitals like that turn out to be model units, which can be treated as benchmarks. We focus not only on the single highest-ranking hospital but also on the group of hospitals which earn the highest ranks [16]. The identification of the weakest performing hospitals is important as well, thus the paper is also focused on such objects.

Stability of rankings has been studied with different vectors of weights. The reference ranking (base ranking) has been constructed using equal weights that relate to the situation in which all 5 criteria are equally important. Thus, the vector of weights for the reference ranking was given by (0.2, 0.2, 0.2, 0.2, 0.2). In the first scenario, we gradually increased the weight of the chosen criterion by 0.1 under the condition that the other criteria have the same relative significance (equal weights). For example, in the case of increasing significance of the first criterion subsequent weight vectors were given by: (0.3, 0.175, 0.175, 0.175, 0.175), (0.4, 0.15, 0.15, 0.15, 0.15); (0.5, 0.125, 0.125, 0.125, 0.125), and in the last iteration: (0.6, 0.1, 0.1, 0.1, 0.1). In this scenario, we used the total of 20 weight vectors plus the initial one (0.2, 0.2, 0.2, 0.2, 0.2). The first scenario reflects the case in which only one criterion is of particular importance for the decision-maker, and it is indifferent to the others, although it still has an impact on the final ranking.

In the second scenario, we assumed that criteria have different relative importance. Rank Order Centroid (ROC) method has been used to derive weights. ROC is one of

the popular methods of deriving weights in case if only the ranking of criteria is given. The advantages of this method are pointed out by Roszkowska [26]. The weights are elicited according to the formula [26]:

$$w_j = \frac{1}{n} \sum_{i=j}^n \frac{1}{r_i} \quad (4)$$

where n – number of criteria, r_i – rank of i th criterion.

In this scenario, we used the total of 120 weight vectors plus the initial one (0.2, 0.2, 0.2, 0.2, 0.2).

In the second approach to the stability analysis, we assessed the stability of rank for each hospital. We used the method of changing weights proposed in [31]. Assuming that all criteria except one have equal importance, we gradually changed the weight of the given criterion, according to the formula:

$$w_j = \begin{cases} 0.2 \left(1 + \frac{i}{1000} \right) & \text{if the weight of } j\text{th criterion has been increased} \\ 0.2 \left(1 - \frac{i}{1000} \right) & \text{if the weight of } j\text{th criterion has been decreased} \end{cases} \quad (5)$$

For the sake of simplicity, we assumed that the weight cannot be lower than 0.01 or higher than 0.99. Then, for each criterion and each hospital, we calculated the range of weights for which the rank of the given hospital does not change.

3.3.2. Similarity measures

In the paper, the overall similarity of rankings is measured using two approaches. In the first one, Spearman's rank correlation coefficients are used. The second approach was proposed by Kukuła in 1989 [17]. It is based on the absolute value of rank differences between two rankings p and q [17]:

$$m_{p,q} = 1 - \frac{2 \sum_{i=1}^n |c_{ip} - c_{iq}|}{n^2 - z} \quad (6)$$

where c_{ip} denotes the rank of i th alternative in ranking p , and c_{iq} denotes the rank of i th alternative in ranking q , while z is given by

$$z = \begin{cases} 0, n \in P \\ 1, n \notin P \end{cases} \quad (7)$$

where P denotes the set of even numbers.

4. Results

4.1. Changes on top positions

The results show that, in general, rankings constructed using different vectors of weights remain similar. Spearman's rank correlation coefficients between rankings built on different weights vectors are presented in Table 1. Coefficients for medium-sized hospitals are presented above the diagonal (white cells) and coefficients for large hospitals are presented below the diagonal (grey cells). In both cases, half of the coefficients are above 0.7. In our case, conclusions from using the approach proposed by Kukuła are similar to these obtained from using a rank correlation coefficient. In this case also half of the measures is at least 0.7 (for large hospitals) and 0.719 (for medium sized-ones).

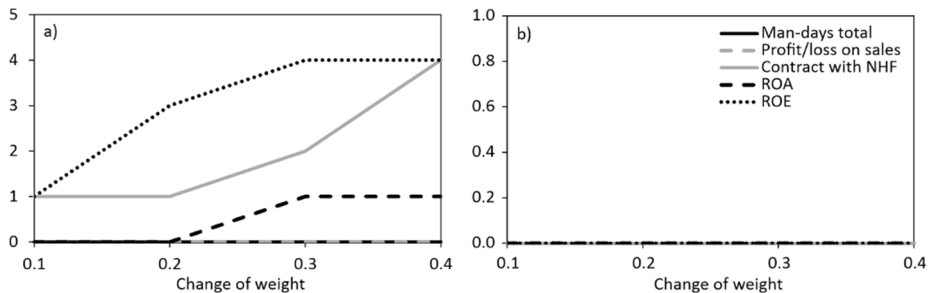


Fig. 2. Change of rank for the highest-ranking large (a) and medium-sized (b) hospitals. Based on the data of district hospitals

In the case of large hospitals, the change in the first position was most clearly seen when the weight of ROE criterion was increasing. After changing this weight by 0.1, the hospital which ranked first in the base ranking fell by 1 rank. After changing the weight by 0.2, the difference amounted to 3 ranks. The increasing importance of the contract with NHF criterion also affected the results very early, but the change of rank was increasing much faster as compared with the previous case. The increasing importance of ROA affected the top position in ranking only if the weight of this criterion was by 0.3 higher than initially. In the case of the medium-sized hospitals, the highest-ranking hospital remained in the first position for all analysed vectors of weights (Fig. 2).

Several highest positions kept is presented in Fig. 3. The entire ranking of large hospitals remained unchanged until the weight of man-days total criterion reached 0.4. Then, only the highest-ranking hospital remained in its previous position. The situation is similar in the case of profit/loss on sales criterion. It should be noted, however, that only 7 highest-ranking hospitals remained on their positions when the weight of this criterion increased to 0.3. Only the highest-ranking hospital remains on its position if the weight of ROA changes by 0.1–0.2. For two other criteria, i.e., ROE, and contract with NHF, the change by 0.1 is sufficient for the change in the first position in the ranking.

Table 1. Spearman's rank correlation coefficients for rankings of medium-sized (in white) and large hospitals (in grey). Iterations 0 (base ranking) – 20

	0	1		4	5		10	11	12	13	14	15	16	17	18	19	20				
0																					
0		0.984*	0.955*	0.99*	0.967*	0.969*	0.919*	0.845*	0.941*	0.903*	0.848*	0.799*	0.788*	0.819*	0.843*	0.482*	0.577*	0.746*	0.678*	0.808*	0.394
1	1.000*		0.932*	0.993*	0.954*	0.961*	0.963*	0.804*	0.968*	0.894*	0.855*	0.871*	0.741*	0.866*	0.835*	0.488*	0.679*	0.698*	0.742*	0.802*	0.398
2	0.973*	0.973*		0.931*	0.939*	0.92*	0.859*	0.955*	0.865*	0.922*	0.789*	0.73*	0.917*	0.721*	0.885*	0.409*	0.493*	0.883*	0.554*	0.86*	0.321
3	0.964*	0.964*	0.936*		0.963*	0.959*	0.943*	0.806*	0.97*	0.901*	0.849*	0.842*	0.744*	0.87*	0.842*	0.485*	0.637*	0.698*	0.747*	0.808*	0.397
4	0.964*	0.964*	0.964*	0.9*		0.938*	0.876*	0.837*	0.918*	0.971*	0.805*	0.749*	0.781*	0.79*	0.935*	0.402*	0.53*	0.735*	0.654*	0.911*	0.303
5	0.955*	0.955*	0.927*	0.973*	0.882*		0.908*	0.803*	0.906*	0.867*	0.943*	0.798*	0.74*	0.791*	0.807*	0.622*	0.584*	0.701*	0.637*	0.771*	0.541*
6	0.855*	0.855*	0.827*	0.818*	0.845*	0.882*		0.734*	0.94*	0.808*	0.846*	0.95*	0.668*	0.87*	0.748*	0.512*	0.814*	0.623*	0.772*	0.717*	0.417
7	0.936*	0.936*	0.973*	0.864*	0.964*	0.855*	0.773*		0.723*	0.86*	0.678*	0.601*	0.99*	0.574*	0.843*	0.34*	0.373*	0.969*	0.401*	0.826*	0.26
8	0.882*	0.882*	0.882*	0.945*	0.800*	0.936*	0.845*	0.773*		0.862*	0.801*	0.878*	0.657*	0.953*	0.805*	0.454*	0.701*	0.606*	0.864*	0.773*	0.369
9	0.964*	0.964*	0.964*	0.900*	1.000*	0.882*	0.845*	0.964*	0.8*		0.715*	0.68*	0.819*	0.735*	0.99*	0.286*	0.466*	0.777*	0.601*	0.979*	0.189
10	0.836*	0.836*	0.791*	0.9*	0.755*	0.945*	0.809*	0.691*	0.855*	0.755*		0.774*	0.616*	0.719*	0.653*	0.802*	0.592*	0.583*	0.572*	0.618*	0.728*
11	0.691*	0.691*	0.691*	0.664*	0.682*	0.727*	0.891*	0.664*	0.755*	0.682*	0.6		0.534*	0.873*	0.624*	0.481*	0.929*	0.489*	0.815*	0.595*	0.396
12	0.891*	0.891*	0.945*	0.791*	0.918*	0.791*	0.718*	0.982*	0.709*	0.918*	0.609*	0.627*		0.503*	0.807*	0.301*	0.317*	0.991*	0.33*	0.796*	0.222
13	0.727*	0.727*	0.718*	0.845*	0.618*	0.827*	0.727*	0.591*	0.945*	0.618*	0.764*	0.745*	0.509*		0.683*	0.433*	0.741*	0.446*	0.964*	0.65*	0.363
14	0.945*	0.945*	0.945*	0.882*	0.991*	0.873*	0.827*	0.955*	0.764*	0.991*	0.773*	0.645*	0.909*	0.573*		0.223*	0.416*	0.766*	0.557*	0.996*	0.128
15	0.691*	0.691*	0.627*	0.791*	0.573*	0.855*	0.736*	0.491*	0.791*	0.573*	0.964*	0.555*	0.409*	0.755*	0.591*		0.347*	0.291*	0.32*	0.185*	0.986*
16	0.391*	0.391*	0.400*	0.355*	0.382*	0.436*	0.673*	0.373*	0.482*	0.382*	0.336*	0.891*	0.382*	0.527*	0.336*	0.364*		0.287*	0.739*	0.399*	0.275
17	0.836*	0.836*	0.900*	0.727*	0.882*	0.727*	0.709*	0.955*	0.664*	0.882*	0.536*	0.636*	0.982*	0.455*	0.873*	0.336*	0.409*		0.275*	0.759*	0.218
18	0.255*	0.255*	0.264*	0.436*	0.091*	0.455*	0.309*	0.109*	0.582*	0.091*	0.509*	0.445*	0.073*	0.745*	0.055*	0.636*	0.5	0.018*		0.531*	0.261
19	0.918*	0.918*	0.918*	0.836*	0.982*	0.827*	0.836*	0.936*	0.727*	0.982*	0.718*	0.664*	0.9*	0.527*	0.991*	0.536*	0.373*	0.882*	-0.009		0.089
20	0.609*	0.609*	0.527*	0.727*	0.464*	0.773*	0.636*	0.373*	0.727*	0.464*	0.9*	0.473*	0.300*	0.718*	0.473*	0.973*	0.355*	0.227*	0.709*	0.418	

Asterisks denote coefficients significant at $\alpha = 0.01$. Based on the data of district hospitals.

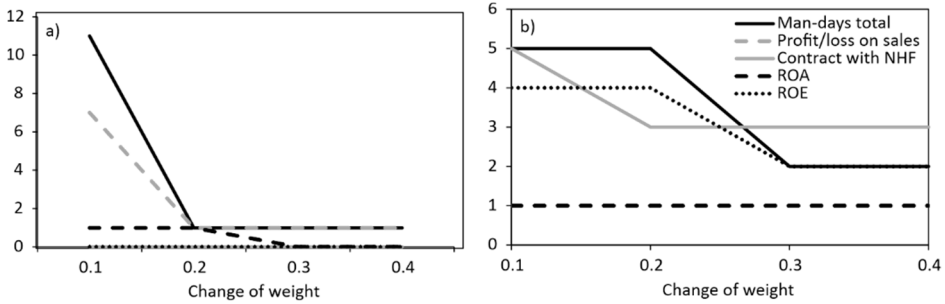


Fig. 3. Number of highest positions kept for large (a) and medium-sized (b) hospitals. Based on the data of district hospitals

In the case of medium-sized hospitals, the situation is slightly different. When the weights of man-days total or the contract with NHF criteria were increased to 0.3, most highest-ranking hospitals (5) maintained their positions. Changes of ROE have a slightly bigger impact as in this case 4 highest-ranking hospitals kept their positions. As the weights of these criteria increased, the number of ranks maintained by highest-ranking hospitals decreased. The decrease started at first in the case of the contract with NHF criterion, and when the weight exceeded 0.4 there was a clear decrease in the case of ROE and man-days total criteria. In the case of ROA criterion, the difference between the newly constructed rankings and the base one occurred very fast, when the weight reached 0.3. Only the highest-ranking hospital kept its position. It kept the first rank in next iterations as well.

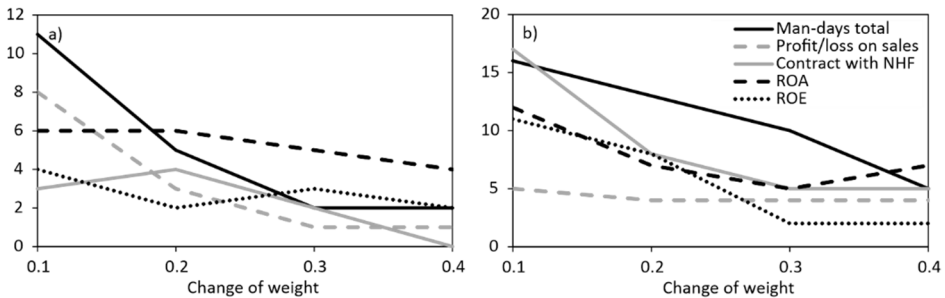


Fig. 4. Number of positions kept for large (a) and medium-sized (b) hospitals. Based on the data of district hospitals

Obviously, the total number of hospitals that kept their positions as the weights were changed was decreasing in both groups (Fig. 4). The clearest decrease can be seen in the case of large hospitals and man-days total and profit/loss on sales criteria. In the case of medium-sized hospitals, the decrease was most explicit in the case of man-days total and contract with NHF criteria. Only 2–4 large hospitals kept their ranks in the case of growing importance of ROE and 4–6 hospitals in the case of the increasing

importance of ROA. For medium-sized hospitals, the impact of profit/loss on sales criterion is relatively the strongest.

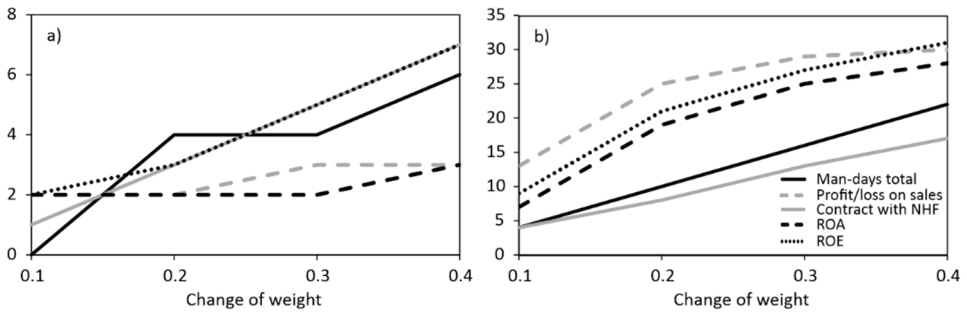


Fig. 5. Maximum rank differences for large (a) and medium-sized (b) hospitals. Based on the data of district hospitals

Maximum rank differences are increasing with the increase of weights (Fig. 5). For large hospitals, differences are the smallest in the case of profit/loss on sales and ROA criteria and for the medium-sized hospitals – in the case of man-days total and contract with NHF criteria. It can be noticed that there were no cases of switching hospitals from the first and last place due to the changes in the relative significance of the criteria.

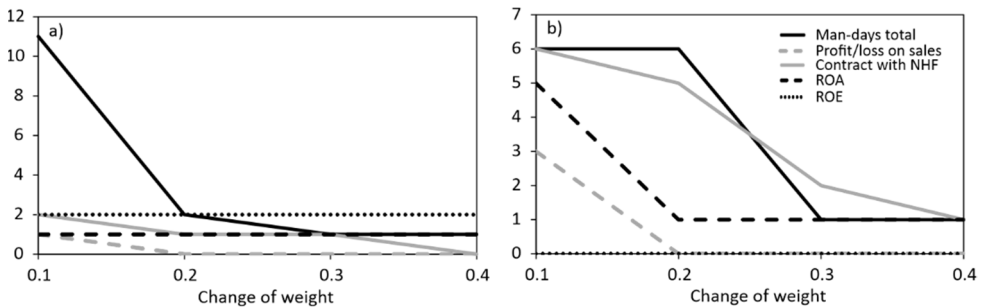


Fig. 6. The number of lowest positions kept for large (a) and medium-sized (b) hospitals. Based on the data of district hospitals

In the case of the preserved number of lowest positions, there are differences between the two groups of hospitals (Fig. 6). For large hospitals and the criteria related to profitability (ROA and ROE), the increase of weights by 0.1 meant that respectively only 1 and 2 hospitals classified in the lowest positions maintained their ranks. The increase of the weight of the man-days total criterion has the slightest effect in the first iteration. 11 hospitals maintained their ranks in this case, which means that the ranking remained unchanged. When the weight of this criterion increased to 0.4 only 2 hospitals ranked lowest maintained their positions. For medium-sized hospitals, there was no change that

would preserve the entire ranking. The most hospitals with lowest ranks (6) maintained their positions when weights of man-days total criterion and contract with NHF criterion increased by 0.1. No hospital maintained its rank when the importance of ROE was increasing.

4.2. Stability of rank for each individual hospital

The range of weights that assure ranks stability for individual hospitals is presented in Tables 2, 3. It can be seen that the majority of large hospitals was most sensitive to changes in ROE and profit/loss of sales criteria (Table 2). Hospitals L1–L4 were least sensitive to changes of man-days total and profit/loss on sales criteria, while hospitals L5–L8 and L11 were least sensitive to changes in ROA criterion. Hospital L10 can be considered an exception because its evaluation was the least sensitive to change in ROE. For 6 hospitals, the maximum lengths of intervals were above 0.8. On the other hand, only in 3 cases, the minimum length of the interval was higher than 0.2.

Table 2. The range weights that assure ranks of stability for large hospitals by criteria

Hospital	Man-days total	Profit/loss on sales	Contract with NHF	ROA	ROE
L1	0.191–0.8304	0.1848–0.99	0.01–0.2156	0.191–0.4492	0.01–0.2042
L2	0.191–0.3134	0.1848–0.312	0.1264–0.2156	0.191–0.2526	0.142–0.2042
L3	0.01–0.3134	0.01–0.312	0.1264–0.2574	0.01–0.2526	0.142–0.317
L4	0.01–0.3052	0.0514–0.4034	0.01–0.2574	0.1046–0.2996	0.114–0.2836
L5	0.01–0.3052	0.1762–0.4034	0.01–0.2668	0.1638–0.5046	0.088–0.218
L6	0.01–0.314	0.1762–0.321	0.01–0.2668	0.1638–0.99	0.0476–0.218
L7	0.01–0.4072	0.01–0.321	0.01–0.4096	0.01–0.99	0.0476–0.3156
L8	0.01–0.3226	0.158–0.2978	0.01–0.2974	0.1326–0.99	0.1438–0.2348
L9	0.01–0.3226	0.158–0.227	0.01–0.2974	0.1326–0.2528	0.1816–0.2348
L10	0.01–0.4272	0.01–0.227	0.01–0.3868	0.01–0.2528	0.1816–0.99
L11	0.01–0.6954	0.01–0.372	0.01–0.5996	0.01–0.99	0.0696–0.99

Based on the data of district hospitals.

In the case of medium-sized hospitals (Table 3), the ranges are smaller, partially due to the greater number of hospitals. Only in the case of 3 hospitals, the length of the interval was greater than 0.2. In general, these hospitals were least sensitive to changes in man-days total and contract with NHF criteria and most sensitive to changes in profit/loss on sales and ROE criteria. There are, however, exceptions. For example, hospital M1 is least sensitive to changes in the relative significance of both profitability criteria.

For the second scenario (ROC weights) (Fig. 7) it is possible to identify both hospitals in the case of which rank is relatively independent of weights, as well as hospitals that are ranked both in the highest and lowest positions for different weight vectors. This finding applies to both large and medium-sized hospitals, even though the overall similarity of rankings remains high.

Table 3. The range of stability for medium-sized hospitals by criteria

Hospital	Man-days total	Profit/loss on sales	Contract with NHF	ROA	ROE
M1	0.01–0.687	0.01–0.9266	0.01–0.6226	0.01–0.99	0.0208–0.99
M2	0.01–0.687	0.01–0.2302	0.01–0.6226	0.01–0.2528	0.1808–0.99
M3	0.01–0.4572	0.0808–0.2302	0.01–0.6072	0.01–0.2528	0.1808–0.4294
M4	0.01–0.4284	0.0808–0.2398	0.01–0.3728	0.0524–0.259	0.1762–0.4294
M5	0.1598–0.4144	0.18–0.252	0.1646–0.3728	0.0524–0.213	0.1684–0.2336
M6	0.1598–0.2984	0.18–0.2568	0.1646–0.2828	0.1336–0.213	0.1668–0.2336
M7	0.01–0.2984	0.1406–0.2368	0.0988–0.2828	0.1534–0.2906	0.1642–0.3508
M8	0.1312–0.304	0.1238–0.22	0.0988–0.313	0.1798–0.2848	0.1618–0.261
M9	0.1598–0.3908	0.1798–0.22	0.064–0.2738	0.1798–0.2106	0.1596–0.261
M10	0.1598–0.2582	0.1798–0.2682	0.064–0.2404	0.1326–0.2106	0.1582–0.347
M11	0.1706–0.2582	0.1744–0.2194	0.1772–0.2404	0.1326–0.2412	0.1564–0.2698
M12	0.1706–0.2182	0.1926–0.2194	0.1772–0.2212	0.0944–0.2324	0.1544–0.2388
M13	0.17–0.2068	0.1966–0.225	0.1672–0.2098	0.0962–0.2148	0.1564–0.2388
M14	0.01–0.2068	0.1966–0.2976	0.01–0.2098	0.118–0.2148	0.1564–0.3398
M15	0.01–0.2936	0.1178–0.298	0.01–0.2962	0.1392–0.3076	0.1474–0.3174
M16	0.01–0.239	0.1748–0.291	0.01–0.253	0.1782–0.335	0.1426–0.26
M17	0.1906–0.239	0.1748–0.2006	0.1958–0.253	0.1994–0.3382	0.157–0.2042
M18	0.1906–0.2468	0.1586–0.2006	0.1958–0.275	0.1994–0.2406	0.1634–0.2042
M19	0.1836–0.2522	0.1832–0.255	0.1866–0.275	0.1948–0.2406	0.1634–0.2024
M20	0.1836–0.2744	0.1832–0.2466	0.1866–0.3458	0.1948–0.3686	0.1358–0.2024
M21	0.085–0.3096	0.1678–0.2466	0.01–0.3602	0.1704–0.2324	0.1398–0.2482
M22	0.085–0.2376	0.1834–0.2544	0.01–0.2472	0.143–0.2156	0.1772–0.2446
M23	0.01–0.2074	0.1866–0.2886	0.01–0.2472	0.1882–0.2156	0.1844–0.2152
M24	0.01–0.2074	0.189–0.2886	0.01–0.2522	0.1882–0.2242	0.1844–0.211
M25	0.01–0.2458	0.189–0.231	0.1188–0.253	0.1426–0.2274	0.1754–0.211
M26	0.1964–0.2458	0.1858–0.2012	0.1896–0.2904	0.1104–0.2024	0.1988–0.2152
M27	0.1964–0.2488	0.1644–0.2012	0.1896–0.2742	0.1598–0.2024	0.1988–0.2578
M28	0.1194–0.2488	0.1336–0.2138	0.01–0.2742	0.1598–0.279	0.187–0.25
M29	0.1194–0.5214	0.1336–0.2352	0.01–0.334	0.1242–0.279	0.165–0.25
M30	0.01–0.5214	0.052–0.214	0.01–0.456	0.1908–0.4574	0.1156–0.2188
M31	0.01–0.5134	0.01–0.214	0.01–0.4282	0.1908–0.329	0.132–0.2188
M32	0.01–0.4568	0.01–0.307	0.01–0.4282	0.1318–0.329	0.1462–0.3168
M33	0.01–0.4568	0.01–0.307	0.01–0.531	0.1822–0.3356	0.1462–0.217
M34	0.01–0.8052	0.01–0.3708	0.01–0.99	0.1822–0.99	0.01–0.217

5. Conclusions

The goal of the paper was to construct rankings of Polish district hospitals using TOPSIS method and to assess the sensitivity of the results to the changes in weights as based on data from 11 large and 34 medium-sized hospitals. To the best of the author's

knowledge, there are no Polish hospital rankings which use well-defined and known multicriteria ranking methods. There are rankings which use algorithms developed by experts, thus this study provides added value in this field. As compared with one of the cited rankings of Polish hospitals [41], in our approach the set of criteria covers more dimensions of hospital performance. The two groups of compared hospitals are homogeneous concerning business conditions (a form of ownership and running emergency ward) and size, which improves the reliability of the results.

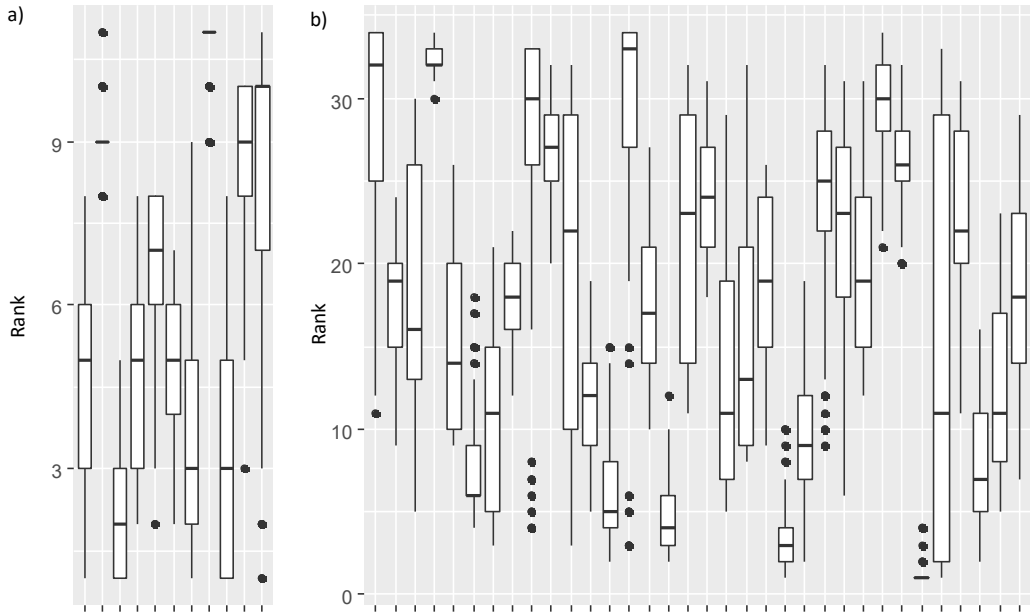


Fig. 7. The distribution of ranks in the analysed groups for large (a) and medium-sized (b) hospitals. Based on the data of district hospitals

The results show that the similarity of rankings is very high if assessed on the overall measure, such as Spearman's rank correlation coefficient, or similarity measure proposed by Kukula [17]. Despite this fact, there are almost no cases in which weight changes do not affect the positions of hospitals. In general, this is not an unambiguously bad signal for hospitals due to the relative nature of the study. The hospitals differ in the criteria selected for the analysis, and the changes in the highest ranks suggest that it is not possible to identify objects which earn the best scores on each criterion. In the case of some hospitals, the ranks turned out to be relatively independent of weights. Those objects may be regarded as some kind of benchmarks but, in general, benchmarking using rankings should be done with care because of the changes in high positions.

We examined the changes both on the highest and lowest positions of rankings. In case of the top positions, it could be seen that the ordering of large hospitals is most

sensitive to changes in importance of ROE and contract with NHF criteria, while rankings of medium-sized hospitals are most sensitive to changes in the weight of ROA. On the other hand, lowest ranks were least prone to change if the importance of profit/loss on sales and ROA criteria (large hospitals) and profit/loss on sales and ROE criteria (medium-sized hospitals) were increasing.

It should be noted that the study has some limitations. Firstly, it does not cover all district hospitals in Poland due to the problem with obtaining data. Secondly, the aspect of quality of care and patients' opinions is not included in the comparison. Broadening the scope of the analysis to cover these topics, as well as identification of characteristic features of most stable and highest-ranking hospitals, are among the most important future directions within this field.

References

- [1] ADELMAN D., *An efficient frontier approach to scoring and ranking hospital performance*, Oper. Res., 2020, Published online in Articles in Advance 17 March 2020, DOI.org/10.1287/opre.2019.1972.
- [2] AFFUL-DADZIE E., NABARESEH S., KOMÍNKOVÁ OPLATKOVÁ Z., *Fuzzy VIKOR Approach. Evaluating quality of internet health information*, [In:] M. Ganzha, L. Maciaszek, M. Paprzycki (Eds.), Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, ACSIS, 2014, 2, 183–190.
- [3] AKCAN S., GÜLDES M., *Integrated multicriteria decision-making methods to solve supplier selection problem. A case study in a hospital*, J. Healthc. Eng., 2019, Article ID 5614892.
- [4] AMINI F., REZAEENOUR J., *Ranking healthcare centers using fuzzy analytic hierarchy process and TOPSIS. Iranian experience*, Int. J. Appl. Oper. Res., 2016, 6 (1), 25–39.
- [5] ARAUJO C., WANKE P., SIQUEIRA M., *A performance analysis of Brazilian public health: TOPSIS and neural networks application*, Int. J. Prod. Perform. Manage., 2018, 67 (9), 1526–1549.
- [6] BRANS J.P., MARESCHAL B., VINCKE P., *How to select and how to rank projects: The PROMETHEE method for MCDM*, Eur. J. Oper. Res., 1986, 24, 228–238.
- [7] BÜYÜKÖZKAN G., ÇİFÇİ G., *A combined fuzzy AHP and fuzzy TOPSIS based strategic analysis of electronic service quality in healthcare industry*, Expert Syst. Appl., 2012, 39 (3), 2341–2354.
- [8] CHEN C.-T., HUNG W.-Z., *Evaluating the service quality of hospital by using TOPSIS with interval type-2 fuzzy sets*, 2017 International Conference on Fuzzy Theory and Its Applications (iFUZZY), Pingtung 2017, 1–5.
- [9] CHOLEWA-WIKTOR M., *Efficiency in public hospitals*, Studia Ekonomiczne Uniwersytetu Ekonomicznego w Katowicach, 2014, 180 (2), 44–54 (in Polish).
- [10] CHURCHMAN C.W., ACKOFF R.L., SMITH Jr. N.M., *An approximate measure of value*, J. Oper. Res. Soc. Am., 1954, 2 (2), 172–187.
- [11] DAVIS P., MILNE B., PARKER K., HIDER P., LAY-YEE R., CUMMING J., GRAHAM P., *Efficiency, effectiveness, equity (E3). Evaluating hospital performance in three dimensions*, Health Policy, 2013, 112, 19–27.
- [12] DU T., RAN L., LI J., CAO X., *Multi-criteria decision-making based on efficiency measurement and empirical study: DEA-TOPSIS*, Chinese J. Manage. Sci., 2017, 25 (7), 153–162.
- [13] HELLWIG Z., *Application of the taxonomic method to the typological distribution of countries according to their level of development and the resources and structure of qualified personnel*, Przegł. Stat., 1968, 4, 307–327 (in Polish).

- [14] HOSSEINI S.M., BAHADORI M., RAADABADI M., RAVANGARD R., *Ranking hospitals based on the disasters preparedness using the TOPSIS technique in Western Iran*, *Hosp. Top.*, 2019, 97 (1), 23–31.
- [15] HWANG C.L., YOON K., *Multiple attribute decision making. Methods and applications*, Springer-Verlag, New York 1981.
- [16] KOŁOCH G., KUSZEWSKI T., SZAPIRO T., *On stability of educational rankings*, [In:] T. Trzaskalik (Ed.), *Multiple criteria decision making '07*, Wydawnictwo AE Katowice, Katowice 2009, 79–101.
- [17] KUKUŁA K., LUTY L., *On the choice of a linear tidying method to assess waste management in Poland in spatial terms*, *Zeszyty Naukowe Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie, Problemy Rolnictwa Światowego*, 2018, 18 (2), 183–192 (in Polish).
- [18] LINGSMA H.F., STEYERBERG E.W., EIJKEMANS M.J.C., DIPPEL D.W.J., SCHOLTE O.P., REAMER W.J.M., VAN HOUWELINGEN H.C., *Comparing and ranking hospitals based on outcome. Results from The Netherlands Stroke Survey*, *QJM-Int. J. Med.*, 2010, 103 (2), 99–108.
- [19] ŁAGOWSKI P., *Financial analysis of key therapeutic entities in the Lower Silesian Voivodship*, [In:] A. Ćwiąkała-Małys, E. Rutkowska-Tomaszewska (Eds.), *Situation of an entrepreneur in post-crisis conditions*, E-Wydawnictwo, Prawnicza i Ekonomiczna Biblioteka Cyfrowa, Faculty of Law, Administration and Economics of the University of Wrocław, Wrocław 2016, 135–149 (in Polish).
- [20] MUTLU M., TUZKAYA G., SENNAROĞLU B., *Multi-criteria decision-making techniques for healthcare service quality evaluation. A literature review*, *Sigma J. Eng. Nat. Sci.*, 2017, 35 (3), 501–512.
- [21] NERMEND K., *Multi-criteria and multidimensional analysis methods in decision support*, Wydawnictwo Naukowe PWN, Warsaw 2017 (in Polish).
- [22] NOJSZEWSKA E., SIELSKA A., GOŁĄB-BELTOWICZ D., *Report on the audit of the financial situation of county hospitals*, Szklarska Poręba 2019, <http://ozpsp.pl/wp-content/uploads/2019/09/Raport-sytuacja-finans.-szpitali-Szklarska-Por%C4%99ba.pdf> (in Polish).
- [23] PRĘDKIEWICZ K., PRĘDKIEWICZ P., WĘGRZYN M., *Profitability of local government hospitals in Poland*, *Nauki Fin.*, 2014, 3 (20), 28–43 (in Polish).
- [24] QI A., PENG K., GAO Y., SHI D., *Using weighted TOPSIS method to evaluate job performance of hospital clinical departments*, *Chinese J. Hosp. Stat.*, 2015, 22 (4), 264–266.
- [25] RAĐENOVIĆ Ž., VESELINOVIĆ I., *Integrated AHP-TOPSIS method for the assessment of health management information systems efficiency*, *Econ. Themes*, 2017, 55 (1), 121–142.
- [26] ROSZKOWSKA E., *Rank ordering criteria weighting methods. A comparative overview*, *Optimum, Studia Ekon.*, 2013, 5 (65), 14–33.
- [27] ROY B., *The outranking approach and the foundations of ELECTRE methods*, [In:] C.A. Bana e Costa (Ed.), *Readings in multiple criteria decision aid*, Springer, Berlin 1990.
- [28] SAATY T., *The analytic hierarchy process*, McGraw-Hill, New York 1980.
- [29] SHAFII M., HOSSEINI S.M., ARAB M., ASGHARIZADEH E., FARZIANPOUR F., *Performance analysis of hospital managers using fuzzy AHP and fuzzy TOPSIS. Iranian experience*, *Glob. J. Health Sci.*, 2016, 8 (2), 137–155.
- [30] SHAFII M., RAFIEI S., ABOOEE F., AMIN BAHRAMI M., NOUHI M., LOTFI F., KHANJANKHANI K., *Assessment of service quality in teaching hospitals of Yazd University of Medical Sciences. Using multi-criteria decision making techniques*, *Osong Publ. Health Res. Persp.*, 2016, 7 (4), 239–247.
- [31] SIELSKA A., *Multicriteria rankings of open-end investment funds and their stability*, *Oper. Res. Dec.*, 2010, 20 (1), 112–129.
- [32] TORKZAD A., BEHESHTINIA M., *Evaluating and prioritizing hospital service quality*, *Int. J. Health Care Qual. Assur.*, 2019, 32 (2), 332–346.
- [33] TRIEMSTRA J.D., POEPELMAN R.S., ARORA V.M., *Correlations between hospitals' social media presence and reputation score and ranking. Cross-sectional analysis*, *J. Med. Int. Res.*, 2018, 20 (11), e28.
- [34] TUZKAYA G., SENNAROĞLU B., KALENDER Z.T., MUTLU M., *Hospital service quality evaluation with IVIF-PROMETHEE and a case study*, *Soc.-Econ. Plan. Sci.*, Elsevier, 2019, 68 (C).

- [35] VÁCHOVÁ L., HAJDÍKOVÁ T., *Evaluation of Czech Hospitals Performance Using MCDM Methods*, [In:] Proceedings of the World Congress on Engineering and Computer Science, WCECS, San Francisco, USA, October 25–27, 2017, Vol. 2.
- [36] WANG C.-H., CHOU H.-L., *Assessment of patient safety management from human factors perspective. A fuzzy TOPSIS approach*, Human Fact. Ergon. Manuf. Serv. Ind., 2015, 25 (5), 614–626.
- [37] WANKE P., AKTAS E., *Efficiency of diabetes treatment*, [In:] C. Kahraman, Y. Topcu (Eds.), *Operations research applications in health care management*, International Series in Operations Research and Management Science, Vol. 262, Springer, Cham 2018, 351–377.
- [38] WIELICKA-GAŃCZARCYK K., *The issue of using economic indicators in assessing the financial condition of medical entities*, Zeszyty Naukowe Politechniki Śląskiej, Seria Organizacja i Zarządzanie, 2015, 78 (1928), 505–516 (in Polish).
- [39] http://www.cmj.org.pl/ranking/wyniki_2019/index.php?typ=zwonk (accessed March 2020).
- [40] <https://rankingi.wprost.pl/szpitala> (accessed March 2020).
- [41] https://www2.deloitte.com/content/dam/Deloitte/pl/Documents/Reports/pl_Ranking_szpitali_publicznych_2017.pdf (accessed March 2020).
- [42] R Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna 2020, URL <https://www.R-project.org/>
- [43] Yazdi M.M., *TOPSIS method for multiple-criteria decision making (MCDM). R package version 1.0*, <https://CRAN.R-project.org/package=topsis> (accessed 20 April 2020).