METHOD FOR SELECTING EXPERT GROUPS AND DETERMINING THE IMPORTANCE OF EXPERTS' JUDGMENTS FOR THE PURPOSE OF MANAGERIAL DECISION-MAKING TASKS IN HEALTH SYSTEM

Ilya Ivlev, Peter Kneppo, Miroslav Barták

Introduction

Current trends in scientific and technological advances are bringing a significant improvements in health care as a result of creation new tools tool to support decision making process for decision-makers (DM) [8], [11], [12], [23], [27]. The experts group approach is used in health sector both clinical [28], [31] and nonclinical [5], [24], [29], [30], [32] decision-making. The goal of this paper is to develop, test and analyse a methodology for determining the qualitative and quantitative composition of an expert group and its application on example of the health technology decision making in the Czech Republic. Health providers face the problem of trying to make decisions in situations where there is insufficient information and also where there is an overload of (often contradictory) information [13]. Regulatory and reimbursement authorities face uncertain choices when considering the adoption of health-care technologies [4]. As the largest share of healthcare expenditure is paid from public sources, the efficient decisions are not only purely technical and financial problem but may be seen as an issue of public interest in the broader terms. This type of decision-making methodology may have a very wide potential also outside the sector of healthcare [14], [16]. Unfortunately the decision-making in the healthcare (and many other part of public sectors) in terms of large investment the Czech Republic is relatively often the object of serious economics as well as legal concerns. To seek for evidence-based methodology that has a potential for evidence-based decision-making is very vital as an opposite to the decision-making influenced by partial individual and group interests.

Many of the most effective models used to find experts are mainly based on the language models [17], [33]. One of the problems with models based on language model frameworks is that they can take into account textual similarities between the query topics and documents [1], [20], [26]. The paper [2] provides two experts finding search strategies modelled to incorporate different types of evidence extracted from the data. The advantage of the modern approaches is machine learning techniques [18], [34] or discriminative probabilistic models [7] a possibility of aggregation of a large number of heterogeneous information. In addition to that, there are the following problems and difficulties: it is difficult to express qualitative information on experts in the quantitative form because information about the candidates varies with time; the experience of a candidate is always varying through time. [8]

In candidate-centric probability estimation approaches for academic expert finding [1], the assessment of an expert is made using generative probabilistic models. In query independent methods [25], knowledge of the expert candidates is presented as a mixture of language models. The person-centric [26] approach is increasingly being used. This paper [19] is based on textual similarities, the author's profile information and the author's citation patterns to try to find academic experts. The methods of Condorcet Fuse [19], Markov chain models [6] and multi-criteria decision-making methods 10] are recognised as representing the most relevant works [3]. The multisensory Data Fusion approach using Dempster-Shafer theory of Evidence together with Shannon's Entropy [20] was used for academic experts finding.

Experts finding is a difficult task because the experts and their skills and knowledge are rare, expensive, constantly changing and varying in depth. When addressing difficult multidisciplinary problems, a combination of knowledge from several experts is often required, especially from experts in various fields. Our method proposal is based on the research of experts' weighting factor determining that reflects the overall competence of the experts when problem solving.

1. Examination of Experts

An examination needs to be performed to determine the weight-coefficients of importance introduced by each criterion of choice. Depending on the scale of the problem, the examination is organised either by a DM in person or by an expert group appointed by the DM. Decisions about the number and competence of experts are made with regard to the scope of the task, the veracity of the evaluations of experts' characteristics and the available resources.

The following tasks need to be solved while creating the expert team: 1) understand the task to be solved by the experts; 2) determine fields of activity linked to the task; 3) decide what share in the team shall be allocated to experts representing each field of activity; 4) determine the number of experts in the team and draft a list thereof; 5) analyse experts' qualifications and edit the draft list of experts; 6) obtain the experts' agreement to work on the team; and 7) finalise the list of experts. Depending on the chosen form of determining experts' preferences, the main requirements of the experts are as follows: 1) competence (reliability and validity of decisions, awareness and reproducible assessment and argumentativeness replies); 2) impartiality; 3) creativity; 4) conformism; 5) team spirit (dependent on quaternary type); 6) relation to the examination; 7) degree of participation in the solving problem; and 8) communication skills (dependent on quaternary type) (Fig. 1).

The experts' characteristics, as listed above, give a comprehensive picture of the qualities that influence the examination results most strongly [20] (properties that are written on a black background are taken into account in our model).

2. Quantity List for the Expert Group

To determine the sufficient number of experts, we needed to find a number H, so that the inequality $W > H$ is true, where W is the dispersive Kendall's coefficient of concordance (coefficient of concordance of experts' opinions. Constant H is selected from the relationship $PW > H = \alpha$ and is fully defined by the level of significance α. It is irrational to select a low level of significance because its decrease results in an increase of H, which, consequently, increases the contingency of the error of the second type [13].

Thereby, we shall determine the necessary quantity of experts that guarantees, at a fixed level of significance, the given critical value of dispersive coefficient of concordance. For reasons of simplification, we shall consider the following relationship true:

$$P\left(W > H\right) \approx P\left[\left(\frac{\chi_{n-1}^2}{n-1}\right) > Hm\right] = \alpha. \tag{1}$$

As $\lim_{n \to \infty} \dfrac{\chi_{n-1}^2}{n-1} = 1$ (n = quantity of parameters), the level of significance of a criterion is determined by the product $H*m$ (m = number of experts), which, even with a small number of experts, can make α sufficiently small.

The survey of experts included recording, in an informative and quantitative form, the experts' opinions about a given problem. The main modes used to survey experts include questionnaires and interviews, discussions and brainstorming. The Delphi method could be used for consensus-building by using a series of questionnaires delivered using multiple iterations [19], [20].

3. Processing of Experts' Evaluations

Processing is needed to obtain generalised data and new information that is concealed in the experts' evaluations. If group evaluations of objects prove doubtful when compared with calculated statistics, it is necessary to determine the reasons for the failed examination. The most typical reasons for failed examinations include the following:
1. Drawbacks in the selection of the expert group: experts' goals did not correspond with the goal of the research (conflict of interests), unsound examination.

Fig. 1: **Expert properties and methods of assessment**
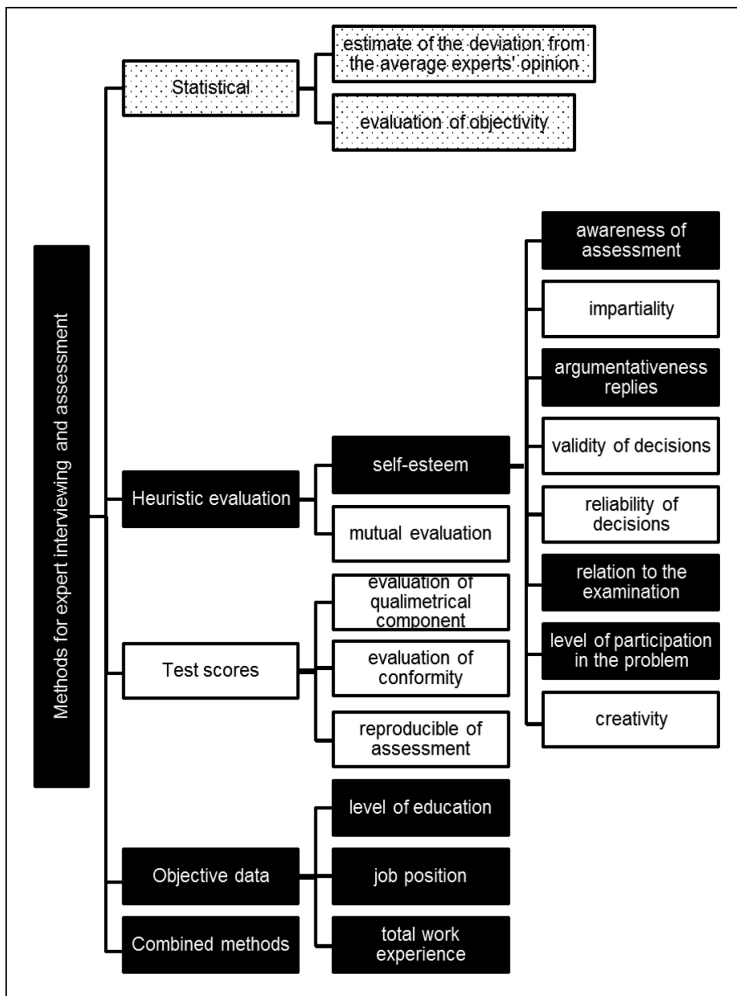
Source: own

2. A conflict of opinion exists. In order to discover different points of view, the experts need to be grouped according to how close their evaluations are. If such grouping proves successful, statistical processing has to be performed for each group separately.

3. Mistakes in the text of the questionnaire, e.g., ambiguous interpretation of questions or use of specific words.

4. Introduction into the questionnaire byway objects. Insufficient conformity of experts'

evaluations does not allow for the group evaluations of all objects to be considered as reliable. If this situation occurs, those evaluations will need to be excluded and the results reprocessed.

Depending on the goals of the expert evaluation, the following main tasks are required to process the survey results: 1) determine experts' competences and generalised evaluations of objects, 2) rank the objects, 3) determine conformity in the experts' opinions,

and 4) determine the relationships among the ranged objects. The following section will describe the individual points relevant to the task we are solving.

3.1 Evaluation of Expert Competence in the Generalised Evaluation of Objects

The first prerequisite to ensure reliability of the examination results is to invite experts who are interested in the results of the examination. Concurrently, the goals of experts have to correspond with the goal of the examination, in general.

It is clear that in present-day conditions, formal indicators of experts (job title, science degree, work experience, number of publications, etc.) can be used only as secondary criteria in identifying experts' total competence that reflects their overall professional skills and qualities. When using the self-evaluation method, the expert provides information about the fields he or she is most competent in [22]. It follows that to evaluate experts' competences; several main methods could be used: self-evaluation, evaluation of a colleague's competence, testing of the experts, and the evaluation of them by the organizers of the examination based on previous examinations. The methods listed above belong to so-called external methods with respect to the conducted examination. However, the method of self-evaluation of competence rates the degree of the expert's self-confidence rather than their real competence. Similarly, during the evaluation method, other people's competence and the group's awareness about each other's abilities play a role. As can be seen, each method has its drawbacks. Other methods of evaluation of experts' competences use posterior data, or the results from the evaluation of objects. Here, experts' competences are evaluated by the degree of conformity of their evaluations to the group evaluation of objects [22]. The essence of this approach lies in the fact that experts who have expressed contradictory opinions receive low grades of competence, and consequently, their evaluations play a less important role when determining the group evaluation. When an expert's evaluation is close to the group's evaluation, the competence of this expert is treated as higher [21] and this fact could be used as a way of experts' competence determining. It should be noted, that grope opinion of experts

with close overall competence level would have higher conformity [27].

3.1.1 Evaluation Based on Objective and Subjective Parameter Assessment

To prevent the results of the self-evaluation method from being a mere reflection of an expert's self-confidence, it is possible to use approaches [9] that provide an objective constituent of knowledge about the expert's competence (Tab. 1).

It is advisable to determine an index of relative self-evaluation by an expert based on the degree of their participation in elaborating the problem as a complex coefficient, which expresses the relationship between the expert and the examination, their participation and their interest. For each question or group of questions on which expert's competence should be evaluated, there is a corresponding scale called the "relative self-evaluation of expert" in the table of expert evaluations.

To prevent the grades in the scale from influencing the self-evaluation, the relative self-evaluation of expert scale contains a list of expert competence properties without any grades.

With this approach, the expert has to underline the properties that, in his or her opinion, determine the level of his or her personal competence. The grades [9] are added by the working group while analysing the collected questionnaires (Tab. 1).

3.1.2 Evaluation of Expert Awareness and His or Her Relevance of Knowledge

Another way (Tab. 2) to determine an expert's weighting factor is via the index of familiarity with the task. It is calculated on the basis of an expert's evaluation of their own familiarity with the problem and indication of typical sources of arguments to support their opinions (index of argumentation, results from summing up the grades in the reference table, index of familiarity with the problem and results from the expert's self-evaluation expressed on a 10-grade scale and multiplied by 0.1 with the purpose of bringing the value to one). In general, the relative self-evaluation of expert index is designed to make the expert perform a self-evaluation of his or her own competence on the given question.

Tab. 1: Questionnaire for evaluation of expert's competence

Objective evaluation								Subjective evaluation	
w_1		w_2		w_3		w_4		w_5	
Job position	Grades	Education	Grades	Total work experience (years)	Grades	Work experience in the problem area	Grades	Level of participation in the problem	Grades
Head of organisation	1.0	Ph.D.	1.0	>10	1.0	>10	1.0	Expert specialises in the given issue	1.0
Deputy head	0.8	Higher education (master)	0.8	10-5	0.8	10-5	0.8	Expert participates in practical work on solving the issue, but the issue does not belong to the expert's indicated specialisation	0.8
Head of department	0.6	Higher education (bachelor)	0.6	<5	0.6	<5	0.6	The issue belongs to the expert's specialisation	0.6
Deputy head of department	0.4	less	0	0	0	0	0	The issue does not belong to the expert's specialisation	0.3

Source: [9]

Tab. 2: Reference table of indices of argumentation ($w_{6,j}$)

$w_{6,j}$	Sources of arguments	Indicators and their weights	Level of source's influence on the expert's opinion			
			I read often and regularly	I read often, but not regularly	I read seldom	I do not read at all
			100%	75%	20%	0%
$w_{6,1}$	Summarising papers by local authors		0.250	0.187	0.050	0
$w_{6,2}$	Summarising papers by foreign authors		0.250	0.187	0.050	0
$w_{6,3}$	Patent information		0.250	0.187	0.050	0
$w_{6,4}$	Companies' reports (catalogues, brochures, recommendations, etc.)		0.250	0.187	0.050	0

Source: own

$$W_6 = \frac{1}{4}\sum_{j=1}^{4} W_{6,j} \qquad (2)$$

Practical research on expert polls show that although self-evaluation methods are not sufficient as the sole criterion for determining expert competence, their application provides a more well-founded selection and evaluation of the experts [9]. To solve the problem of setting the weight coefficients of experts and, thus, determining the probability of obtaining reliable evaluations, we propose to create a comprehensive method based on the approach that combines various competence evaluation methodologies: a self-evaluation of experts about their own competence on the problem; an introduction of grades for objective data and an appreciation of the index of argumentation. Thereby, the competence index of an expert can be treated as a probability of the expert's giving a reliable evaluation, where $0 \le W_e \le 1$.

4. Results and Discussion
4.1 Determining the Quality of the Expert Group

The methodology was used for the selection of experts for the propose of rational selection of large medical equipment such as, computed tomography (CT), mammographic digital X-ray systems (MAM), magnetic resonance imaging (MRI), radiographic/fluoroscopic systems (general purpose) (RFS), ultrasonic scanning systems (cardiac) (USC) and ultrasonic scanning systems (general purpose) (USG). The list of most health care facilities or departments in the Czech Republic where CT, MAM, MRI, RFS, USC and USG are located was formed. Potential experts were asked to answer the web-based questionnaire. The purpose of experts groups' creation was presented in each questionnaire. Candidates for the experts were managers (heads of departments, heads of clinics, etc.) and employees (physicians, biomedical engineers, radiological assistants, senior technicians, etc.) from the departments of radiology, imaging methods, radio diagnostics, interventional radiology, departments of cardiology and the Institute of radiology and the other departments in hospitals of all levels of organisation and all regions of the Czech Republic. The selection of experts was conducted with 872 employees from 422 health facilities in the Czech Republic (Tab. 3).

The method of selecting the most knowledgeable experts for the task of selecting and procuring medical equipment for hospitals is based on 1) the expert's overall work experience, 2) experience in solving tasks, 3) level of education and scientific record, 4) interest in solving the particular task, 5) current position and 6) awareness of how to solve the task. This study also considered the

Tab. 3: The participants of the survey

The task	The number of health care facilities' staffed potential experts	The number of questionnaires sent	The number of responses
Selection of MRI	34	60	19 (31.7%)
Selection of mammographic digital X-ray systems	68	125	18 (14.4%)
Selection of USC	101	190	22 (11.6%)
Selection of CT	89	162	15 (9.3%)
Selection of USG	116	116	9 (7.0%)
Selection of RFS	14	219	13 (6.0%)
Total	422	872	96 (11.0%)

Source: own calculations

7) relevance of the expert's knowledge and 8) the overall self-evaluation concerning their total competence in solving the task.

The data obtained from the questionnaire of experts did not follow a normal (Shapiro-Wilk test) distribution ($p < 0.001$) ("Source of argumentation" [$p = 0.029$] and "Self-ranking" [$p = 0.003$]). An analysis of the results revealed (Mann-Whitney U-test) that there was no reason to reject the null hypothesis that the inter-group values of the compared (total rating w_{TR} and self-rating w_{SR}) characteristics were homogeneous ($p = 0.285$). A medium-strength positive correlation (Spearman's rank correlation) was found between the two measures ($r = 0.550$, $p < 0.001$) (Fig. 5).

Examinations of the correlations between the components that determine the total weight of the experts were performed. The correlations between "Work experience in the problem area" and "Education" (r = 0.268; p = 0.01) and "Work experience in the problem area" and "Job position" (r = 0.342; p = 0.001) suggest that solving the task of selecting large medical equipment participating potential experts with higher level of education and higher operating positions. The task of selecting large medical equipment is categorised as a managerial task should be solved by the most experienced professionals filling a managerial role.

The moderate correlation between "*Work experience in the problem area*" and "Sources of arguments" ($r = 0.394$, $p < 0.001$) and the weak correlation between "Work experience in the problem area" and "*Level of participation in the problem*" ($r = 0.239$, $p < 0.022$), as well as "Sources of arguments" and "Job position" ($r = 0.231$, $p = 0.027$), indicate that higher levels of expert experience equate to higher levels of participation in solving the task, higher levels of theoretical preparation and higher levels of overall competence.

The w_{SR} index was obtained as a result of experts' self-rankings on a scale from zero (*I am NOT competent in addressing the issue of selection*) to 10 (*I am competent in addressing the issue of selection*). The w_{SR} most closely correlated with "Work experience in the problem area" ($r = 0.519$, $p < 0.001$), that is, 52% of the general weighting factor consists of the experts' «Work experience in the problem area». The next parameters most closely correlated to the w_{SR} were the experts' levels of argumentativeness (theoretical preparation, source of arguments and awareness) ($r = 0.440$, $p < 0.001$), "Job position" ($r = 0.319$, $p = 0.002$) and "Education" ($r = 0.280$, $p = 0.007$). The presence of the above correlations indicates that the index of experts' self-rankings is dependent on "Work experience in the problem area", "Sources of arguments", "Job position" and "Education". A statistically significant association was not detected between w_{ST} and the experts' total work experience ($p = 0.089$) or w_{SR} and level of experts' participation in the problem solving task of selecting medical equipment ($p = 0.200$).

Fig. 2 shows that the least sensitive (less varying, depending on the expert) indicators were "Total work experience", "Education", "Work experience in the problem area" and "Level of participation in the problem". These results are indicative of the fact that specialists with a high level of education and high-quality positions, whose work was related to the problem of selection, were pre-selected as experts.

Fig. 2: The list of experts in order by their comprehensive assessment of competence (values obtained from the evaluation of experts on the objective and subjective criteria and evaluation of their own opinions about their own competence)
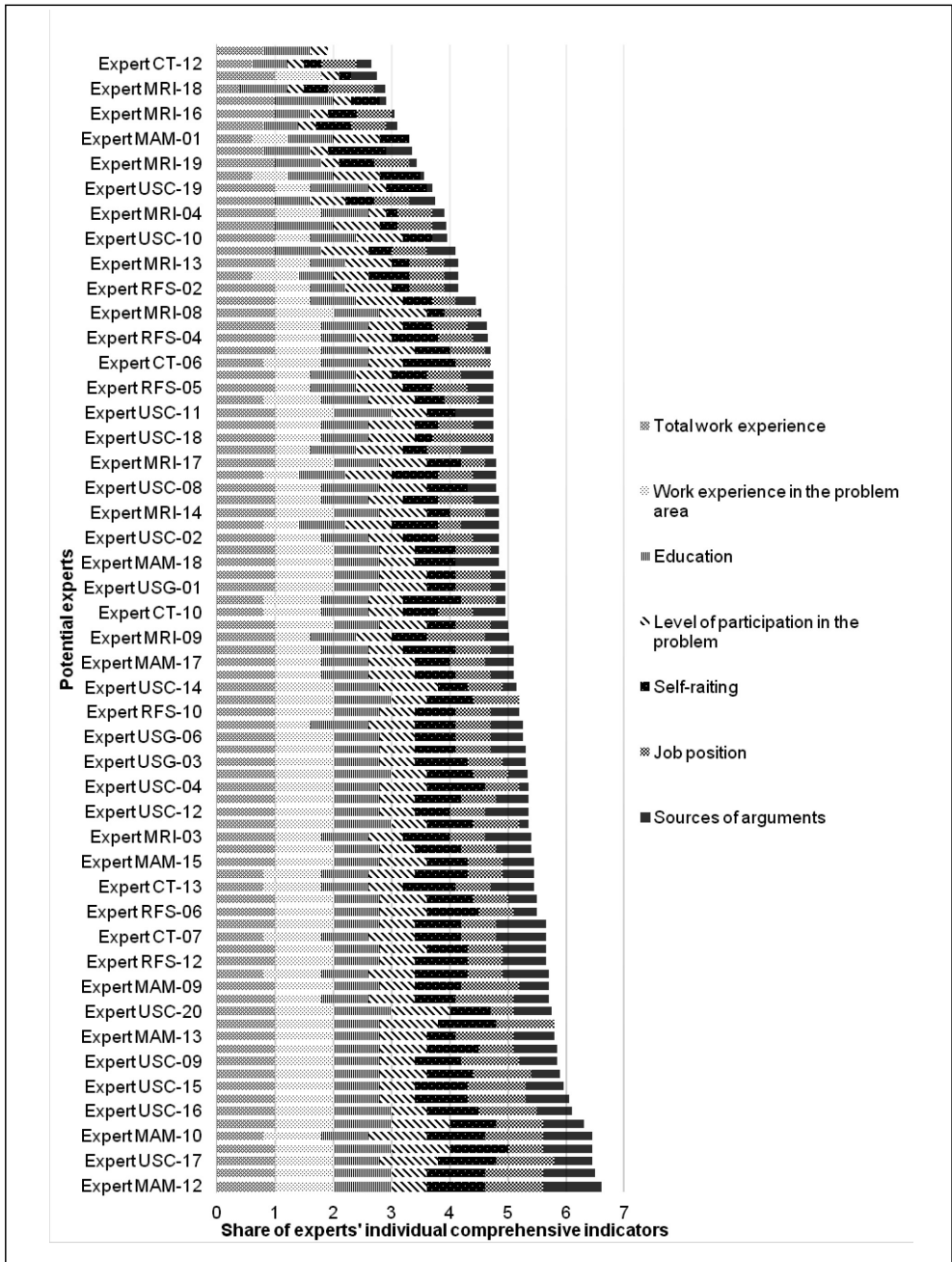
Source: own calculations
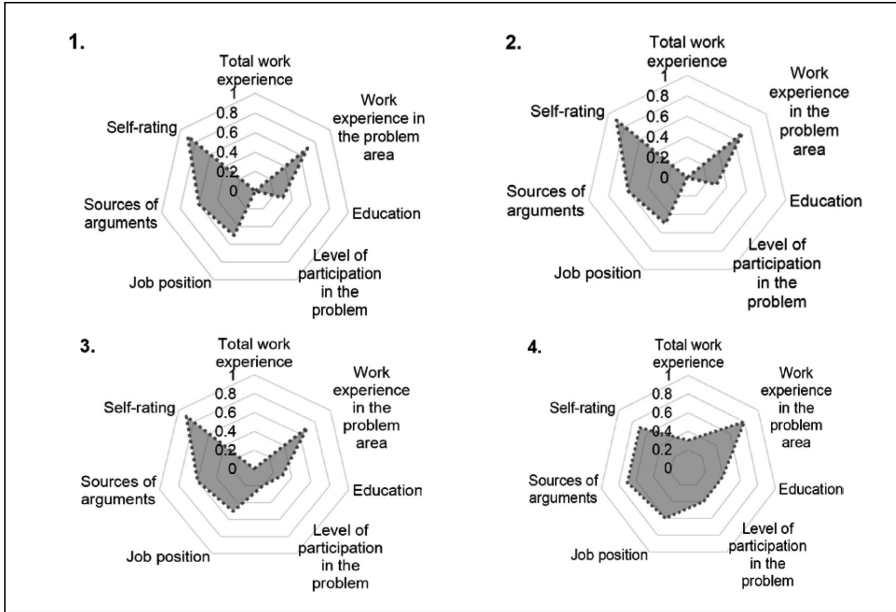
Source: own calculations

4.2 The Overall Weight of Experts' Competence

Four different calculation models (Fig. 3) for the total competence weighting factor were investigated to determine the final model for calculating the total weight of competence of each expert.

In the first method:

$$W_e^{(1)} = \frac{1}{2}\left(w_{ST} + w_{TR}\right), \tag{3}$$

where

$$w_{TR} = \frac{1}{n}\sum_{i=1}^{n} w_i, \tag{4}$$

the weight of each of the coefficients (w_{ST} and w_{TR} [formula 4]) are considered equal, i.e., the contribution of each of the coefficients to the total weight $W_e^{(1)}$ is the same (as in $W_e^{(2)}$ and $W_e^{(3)}$) (Fig. 5). Since the arithmetic mean is not a robust statistic (is subject to strong influence of large deviations), $W_e^{(1)}$ increasingly relies on the uncertainty of w_{ST} (Fig. 6-B).

The second method of calculation

$$W_e^{(2)} = \sqrt{w_{ST}^2 + w_{TR}^2}, \tag{5}$$

is an adaptation of the calculation of the A-type uncertainty measurement in a calibration, where the input values are correlated, as in this case ($r = 0.55$, $p < 0.001$) (Fig. 5). The w_{TR} is viewed as uncertainty in the assessment of competence obtained on the basis of objective and subjective parameters. The w_{ST} is viewed as the uncertainty contributed by other unaccounted factors. The experts' weighting factors obtained by this method were very accurate reproductions of the estimate $W_e^{(1)}$ ($r = 0.998$, $p < 0.001$) (Fig. 6-B, Fig. 6-C) and faithfully reproduced the estimates for w_{ST} and w_{TR}.
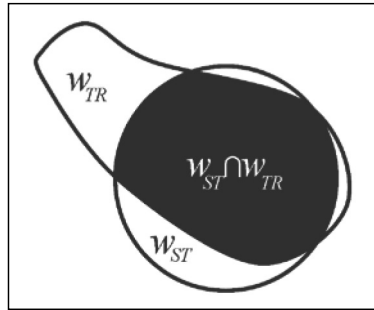
In the third model,

$$W_e^{(3)} = w_{ST} \times w_{TR}, \qquad (6)$$

the single estimates w_{ST} and w_{TR} were viewed as the probability (formula 7) of the simultaneous occurrence of two independent events.

$$P(AB) = P(A) \times P(B). \qquad (7)$$

Hence, the general probability $W_e^{(3)}$ was regarded as the value contained in the region of overlap of probabilities w_{ST} and w_{TR} (Fig. 4).

Fig. 4: The overlap of probabilities w_{ST} and w_{TR}

Source: own

With the use of the product, both variables w_{ST} and w_{TR} were endowed with the property of equal importance (equivalent in the same manner as in $W_e^{(1)}$ and $W_e^{(2)}$).

A very strong correlation was found between the weighting factors $W_e^{(1)} - W_e^{(2)}$ ($r = 0.998$, $p < 0.001$) and $W_e^{(1)} - W_e^{(3)}$ ($r = 0.998$, $p < 0.001$). From the figure (Fig. 6-A, Fig. 6-B, Fig. 6-C), it seems that the nature of the curve $W_e^{(k)}$ corresponded more to w_{ST} and less with w_{TR} (Fig. 5). When using multiplication during the $W_e^{(3)}$ calculations, the slope of the curve changes (Fig. 6-D).

The slope of the curve changes, because of the fact (due to the multiplication) that experts with larger w_{ST} and w_{TR} get more weight $W_e^{(3)}$ and experts with lower values of w_{ST} and w_{TR} get less weight $W_e^{(3)}$, in the context of this task is no longer justified. This is considered artificial expert dilution and is leading to a loss of proportional relationships between them.

The fourth approach to the definition of W_e is fundamentally different:

$$W_e^{(4)} = w_{TR}^+ = \frac{1}{n+1}\left(\sum_{i=1}^{n} w_i + w_{ST}\right). \qquad (8)$$

The w_{ST} here was part of $W_e^{(4)}$, along with other components (w_i). The estimate $W_e^{(4)}$ corresponded more to w_{TR} ($r = 0.966$, $p < 0.001$) and less to w_{ST} ($r = 0.714$, $p < 0.001$) (Fig. 5), which was contradictory to $W_e^{(1)}$, $W_e^{(2)}$ and $W_e^{(3)}$. Thus, this was also true when the w_{ST} became a subsidiary corrective measure of $W_e^{(4)}$. The shape of the curve $W_e^{(4)}$ (Fig. 6-A) was more consistent with the comprehensive assessment w_{TR}, which was calculated on the basis of objective and subjective information.

The use of the model $W_e^{(4)}$ eliminated violations of proportionality extreme values, and considerably reduced the effect of uncertainty w_{ST}. This effect is desirable because the importance attached to w_{ST} is unjustified. The w_{ST} depends too much on experts' own psychological states and their understanding of the grading scale. The model for calculating $W_e^{(4)}$ is the most suitable for the determination of experts' general competences. The estimate $W_e^{(4)}$, for obvious reasons, less effectively reproduced ($r = 0.690$, $p < 0.01$) the union of following weight functions, w_{ST} and w_{TR}, than the estimates obtained by other methods ($W_e^{(1)}$, $W_e^{(2)}$ and $W_e^{(3)}$). (Fig. 5).
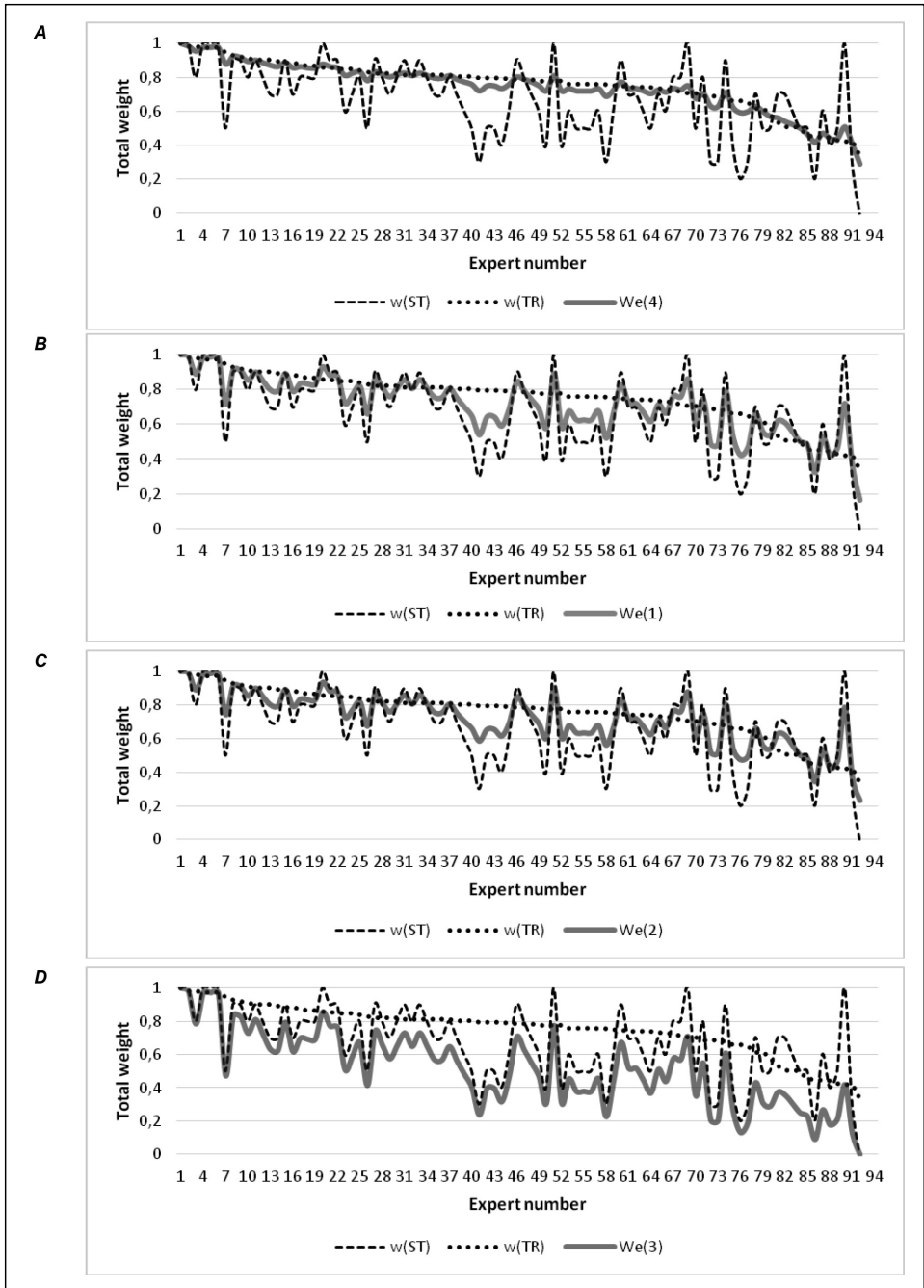
Fig. 5: The correlation between the estimates 1) obtained by the self-assessment w_{ST} and 2) obtained on the basis of objective and subjective data w_{TR}
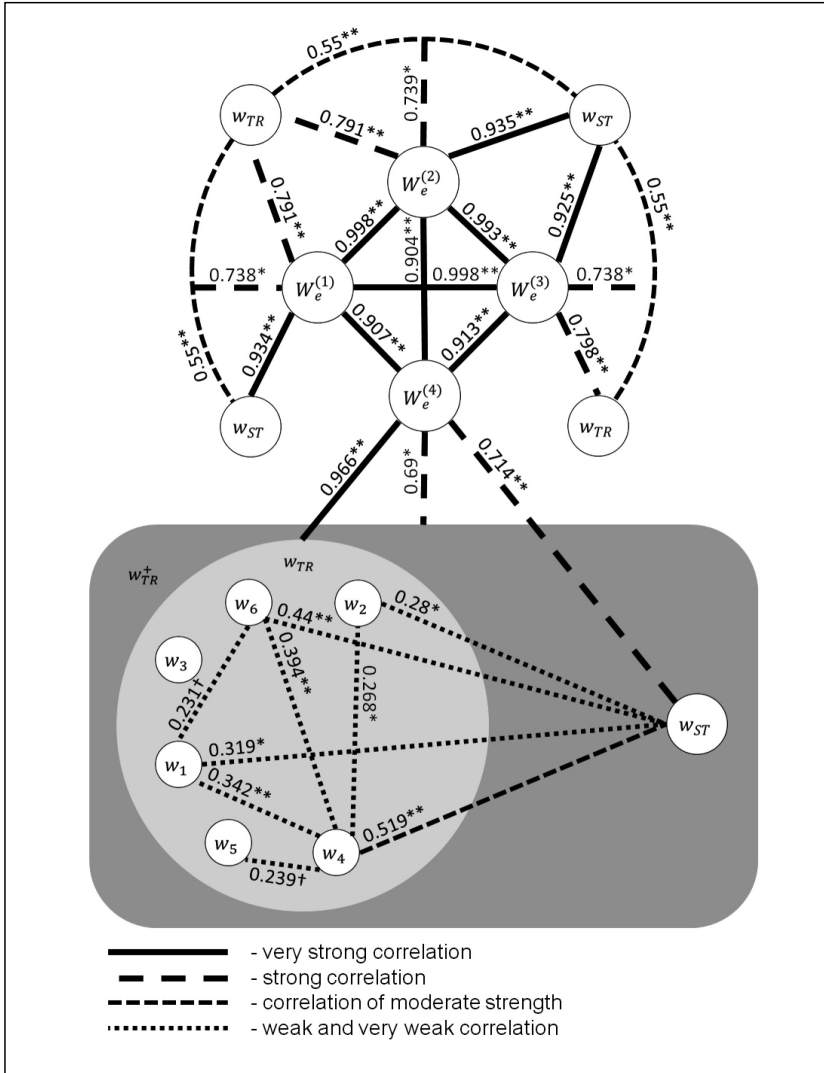
Source: own calculations

† - $p \leq 0.05$; * - $p \leq 0.01$; ** - $p \leq 0.001$

Fig. 6: Comparison of the curves w_{ST} and w_{TR} with $A - W_e^{(4)}$, $B - W_e^{(1)}$, $C - W_e^{(2)}$ and $D - W_e^{(3)}$

Source: own calculations

The greatest difficulties in establishing an expert group include the following: 1) the complexity of taking into account the diverse properties of the expert, 2) the integration of humans' psycho-physiological characteristics (a tendency to take risks, a tendency to formalise and subconscious preference for various numbers), 3) the complexity of describing of the study area and 4) accounting for all components.

Conclusions

This paper developed, tested and analysed a method for determining the qualitative and quantitative composition of an expert group. As a result of the application procedure, each potential expert was evaluated on eight complex-valued criteria based on objective and subjective data. After the evaluation, each expert was given their own weighting factor regarding the importance of their judgments. The most important expert properties that might be considered for determination of their general weight included the following: 1) professional competence (reliability and validity of the decisions rendered), 2) impartiality, 3) objectivity, 4) concern in participating in the examination, 5) ability to operate on a scale of relations and a scale of probability, and 6) ability to take into account the large number of gradation scales. Another important factor is the reproducibility of the results, which can be assessed by numerous questionnaires. Thus, the accuracy of judgment of correctly formed expert groups is sufficiently large, and the error does not exceed 5–10%. This method can be used in the formation of an expert group for virtually any application. This approach to the decision-making processes in the health sector may be understood as a contribution to the evidence based health policy with respect both to the nonclinical as well as clinical decisions. This methodology may be a partial contribution in some fields of scientific and technological forecasting, managerial decision making, quality assessment and operational research both in public and private sector.

References

[1] BALOG, K., AZZOPARDI, L. and DE RIJKE, M. Formal models for expert finding in enterprise corpora. In: *Proceedings of the Twenty-Ninth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. New York: ACM Press, 2006. pp. 43-50. ISBN 1595933697. DOI: 10.1145/1148170.1148181.

[2] BALOG, K., AZZOPARDI, L. and DE RIJKE, M. A language modeling framework for expert finding. *Information Processing and Management*. 2009, Vol. 45, Iss. 1, pp. 1-19. ISSN 0306-4573. DOI: 10.1016/j.ipm.2008.06.003.

[3] BRAUERS, W., ZAVADSKAS, E. and PLANNING, T. A Multi-Objective Decision Support System for Project Selection with an Application for the Tunisian Textile Industry. *E+M Ekonomie a Management*. 2012, Vol. 15, Iss. 1, pp. 28-43. ISSN 1212-3609.

[4] CLAXTON, K., SCULPHER, M. and DRUMMOND, M. A rational framework for decision making by the National Institute For Clinical Excellence (NICE). *Lancet*. 2002, Vol. 360, Iss. 9334, pp. 711-715. ISSN 0140-6736. DOI: 10.1016/S0140-6736(02)09832-X.

[5] DIONNE, F., MITTON, C., MACDONALD, T., MILLER, C. and BRENNAN, M. The challenge of obtaining information necessary for multi-criteria decision analysis implementation: the case of physiotherapy services in Canada. *Cost effectiveness and resource allocation.* 2013, Vol. 11, Iss. 1, pp. 1-16. ISSN 1478-7547. DOI: 10.1186/1478-7547-11-11.

[6] DWORK, C., KUMAR, R., NAOR, M. and SIVAKUMAR, D. Rank aggregation methods for the Web. In: *Proceedings of the tenth international conference on World Wide Web - WWW '01*. New York: ACM Press, 2001. pp. 613-622. WWW '01. ISBN 1581133480.

[7] FANG, Y., SI, L. and MATHUR, A. Discriminative models of integrating document evidence and document-candidate associations for expert search. In: *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval - SIGIR '10*. New York: ACM Press, 2010. pp. 683-690. ISBN 978-1450301534.

[8] FERREYRA RAMÍREZ, E.F.F. and CALIL, S.J. Connectionist Model to Help the Evaluation of Medical Equipment Purchasing Proposals. *World Congress on Medical Physics and Biomedical Engineering 2006*. 2007, Vol. 14, pp. 3786-3789. ISSN 1680-0737.

[9] GOLUBKOV, E. *Marketing Research: Theory, Methodology and Practice*. Moscow: Finpress, 2008. ISBN 978-5-8001-0093-8.

[10] IVLEV, I., KNEPPO, P. and BARTÁK, M. Multicriteria Decision Analysis: a Multifaceted Approach to Medical Equipment Management. *Technological and Economic Development of Economy*. 2014, Vol. 20, Iss. 3, pp. 576-589. ISSN 2029-4913. DOI: 10.3846/20294913.2014.943333.

[11] IVLEV, I., BARTÁK, M., KNEPPO, P. Methodology for selecting experts groups for the purpose of decision-making tasks. *Value in Health.* 2014, Vol. 17, No. 7, pp. A580-A580. ISSN 1098-3015. DOI: 10.1016/j.jval.2014.08.1961.

[12] IVLEV, I., VACEK, J. and KNEPPO, P. Multi-criteria decision analysis for supporting the selection of medical devices under uncertainty. *European Journal of Operational Research*. 2015 (Forthcoming). ISSN 0377-2217.

[13] JONES, J. and HUNTER, D. Qualitative Research: Consensus methods for medical and health services research. *BMJ*. 1995, Vol. 311, Iss. 7001, pp. 376-380. ISSN 0959-8138. DOI: 10.1136/bmj.311.7001.376.

[14] KAKLAUSKAS, A, ZAVADSKAS, E. and ŠAPARAUSKAS, J. Knowledge management and decision making. *Ukio Technologinis ir Ekonominis Vystymas*. 2004, Vol. 10, Iss. 4, pp. 142-149. ISSN 1822-3613. DOI: 10.1080/13928619.2004.9637671.

[15] KENDALL, M. and GIBBONS, J.D. *Rank Correlation Methods*. 5th ed. London: A Charles Griffin Book, 1990. ISBN 0852643055.

[16] LIBERATORE, M.J. and NYDICK, R.L. The analytic hierarchy process in medical and health care decision making: A literature review. *European Journal of Operational Research*. 2008, Vol. 189, Iss. 1, pp. 194-207. ISSN 0377-2217. DOI: 10.1016/j.ejor.2007.05.001.

[17] LIU, P. An approach to group decision making based on 2-dimension uncertain linguistic information. *Technological and Economic Development of Economy*. 2012, Vol. 18, Iss. 3, pp. 424-437. ISSN 2029-4913. DOI: 10.3846/20294913.2012.702139.

[18] MACDONALD, C. and OUNIS, I. Learning Models for Ranking Aggregates. In: *Advances in Information Retrieval*. Dublin: Springer Berlin Heidelberg, 2011. pp. 517-529. ISBN 978-3-642-20160-8.

[19] MONTAGUE, M. and ASLAM, J.A. Condorcet fusion for improved retrieval. In: KALPAKIS, K., GOHARIAN, N. and GROSSMAN, D. (Eds.). *Proceedings of the eleventh international conference on Information and knowledge management - CIKM '02*. New York: ACM Press, 2002. pp. 538-548. ISBN 1581134924.

[20] MOREIRA, C. and WICHERT, A. Finding academic experts on a multisensor approach using Shannon's entropy. *Expert Systems with Applications*. 2013, Vol. 40, Iss. 14, pp. 5740-5754. ISSN 0957-4174. DOI: 10.1016/j.eswa.2013.04.001.

[21] ORLOV, A. *Teoriya prinyatiya resheniy [Theory of decision-making]*. Moscow: Ekzamen, 2006. ISBN 5-472-01393-3.

[22] PAVLOV, A. and SOKOLOV, B. *Metody obrabotki ekspertnoj informacii [Methods of Experts' Information Processing]*. Saint Petersburg: Saint Petersburg State University of Aerospace Instrumentation, 2005.

[23] PECCHIA, L. and BATH, P.A. AHP and risk management: a case study for assessing risk factors for falls in community-dwelling older patients. In: *Proceedings of the International Symposium on the Analytic Hierarchy Process 2009*. 2009, pp. 1-15.

[24] PECCHIA, L., et al. User needs elicitation via analytic hierarchy process (AHP). A case study on a Computed Tomography (CT) scanner. *BMC medical informatics and decision making*. 2013, Vol. 13, Iss. 1, pp. 1-11. ISSN 1472-6947. DOI: 10.1186/1472-6947-13-2.

[25] PETKOVA, D. and CROFT, W. Hierarchical Language Models for Expert Finding in Enterprise Corpora. In: *18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'06)*. Arlington, VA: IEEE, 2006. pp. 599-608. ISBN 0-7695-2728-0.

[26] SERDYUKOV, P. and HIEMSTRA, D. Modeling documents as mixtures of persons for expert finding. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Glasgow: Springer Berlin Heidelberg, 2008. pp. 309-320. ISBN 3540786457.

[27] SORENSON, C., DRUMMOND, M. and BHUIYAN KHAN, B. Medical technology as a key driver of rising health expenditure: disentangling the relationship. *ClinicoEconomics and outcomes research* [online]. 2013, Vol. 5 [cit. 2014-06-24], pp. 223–34. ISSN 1178-6981. DOI: 10.2147/CEOR.S39634.

[28] SUNER, A., ÇELİKOĞLU, C.C., DICLE, O. and SÖKMEN, S. Sequential decision tree

using the analytic hierarchy process for decision support in rectal cancer. *Artificial intelligence in medicine*. 2012, Vol. 56, Iss. 1, pp. 59-68. ISSN 1873-2860. DOI: 10.1016/j.artmed.2012.05.003.
[29] ŠOLTÉS, V., GAVUROVÁ, B. The functionality comparison of the health care systems by the analytical hierarchy process method. *E+M Ekonomie a Management.* 2014, Vol. 17, Iss. 3, pp. 100-117. ISSN 1212-3609. DOI: 10.15240/tul/001/2014-3-009.
[30] VELMURUGAN, R. and SELVAMUTHUKUMAR, S. The analytic network process for the pharmaceutical sector: Multi criteria decision making to select the suitable method for the preparation of nanoparticles. *Daru : journal of Faculty of Pharmacy, Tehran University of Medical Sciences*. 2012, Vol. 20, Iss. 1, pp. 1-14. ISSN 1560-8115. DOI: 10.1186/2008-2231-20-59.
[31] VENHORST, K., ZELLE, S.G, TROMP, N. and LAUER, J.A. Multi-criteria decision analysis of breast cancer control in low- and middle-income countries: development of a rating tool for policy makers. *Cost effectiveness and resource allocation.* 2014, Vol. 12, pp. 13. ISSN 1478-7547. DOI: 10.1186/1478-7547-12-13.
[32] VINCENT, C.J., LI, Y. and BLANDFORD, A. Integration of human factors and ergonomics during medical device design and development: It's all about communication. *Applied Ergonomics*. 2014, Vol. 45, No. 3, pp. 413-419. ISSN 1872-9126. DOI: 10.1016/j.apergo.2013.05.009.

[33] WEI, G. and ZHAO, X. Methods for probabilistic decision making with linguistic information. *Technological and Economic Development of Economy*. 2014, Vol. 1, Iss. 20, pp. 1-17. ISSN 2029-4913. DOI: 10.3846/20294913.2014.869515
[34] YANG, Z., TANG, J., WANG, B., GUO, J. and LI, J. Expert2Bólè : From Expert Finding to Bólè Search. In: *Proceedings of the 15th ACM conference on knowledge discovery and data mining (KDD'09)*. Paris, 2009. pp. 1-4. ISBN 9781605584959.

Ing. Ilya Ivlev, Ph.D.
Czech Technical University in Prague
Faculty of Biomedical Engineering
Department of Biomedical Technology
ilya.ivlev@fbmi.cvut.cz

prof. Ing. Peter Kneppo, DrSc.
Czech Technical University in Prague
Faculty of Biomedical Engineering
Department of Biomedical Technology
kneppo@fbmi.cvut.cz

PhDr. Miroslav Barták, Ph.D.
Jan Evangelista Purkyne University
Faculty of Social and Economic Studies
Department of Social Work
miroslav.bartak@ujep.cz

Abstract

METHOD FOR SELECTING EXPERT GROUPS AND DETERMINING THE IMPORTANCE OF EXPERTS' JUDGMENTS FOR THE PURPOSE OF MANAGERIAL DECISION-MAKING TASKS IN HEALTH SYSTEM

Ilya Ivlev, Peter Kneppo, Miroslav Barták

This work aims to develop a methodology for determining the qualitative composition of an expert group and the weighting factor regarding the importance of expert's judgments for the purpose of participating in decision-making. It is based on the expert's overall work experience, experience in solving tasks, level of education and scientific record, interest in solving the particular task, current position and awareness of how to solve the task. This study also considered the relevance of the expert's knowledge and the overall self-evaluation concerning their total competence in solving the task. For the purpose of validating the methodology, 96 potential experts (physicians, biomedical engineers, radiological assistants, medical physicists, etc.) from 72 health facilities in the Czech Republic were interviewed through a web-based questionnaire. The calculation model that was selected was able to eliminate errors in estimating the proportionality of extreme values and reduces the impact of uncertainty in the experts' overall self-evaluations concerning their total competence. A statistically significant correlation was found between the complex weighting factor and the following characteristics: the expert's experience in dealing with similar tasks ($r = 0.512$, $p < 0.001$), the expert's theoretical background (awareness) and the relevance of the expert's knowledge ($r = 0.440$, $p < 0.001$), the expert's current position ($r = 0.319$, $p = 0.002$) and the level of his or her education and scientific record ($r = 0.280$, $p = 0.007$). The developed methodology may be especially useful in scientific and technological forecasting, medical and managerial decision-making, quality assessment and operational research.