

# REVISION AND EVALUATION OF ADMISSION PROCESS FOR MASTER'S DEGREE PROGRAMME IN DIETETICS AND FOOD SCIENCE AT MEDICAL UNIVERSITY

M. Panczyk<sup>1</sup>, B. Sińska<sup>2</sup>, J. Belowska<sup>1</sup>, A. Zarzeka<sup>1</sup>, J. Gotlib<sup>1</sup>

<sup>1</sup> *Division of Teaching and Outcomes of Education, Faculty of Health Sciences, Medical University of Warsaw (POLAND)*

<sup>2</sup> *Division of Human Nutrition, Faculty of Health Sciences, Medical University of Warsaw (POLAND)*

## Abstract

### Introduction:

Despite the numerous scientific papers concerning shaping admission policy on medical universities, there is still no consensus in the subject of efficiency and reliability of tools for selecting candidates to universities for faculties connected with educating in the field of the selected specialisations in health sciences. Between 2012 and 2013, the Medical University of Warsaw (MUW) made certain changes in entrance test content outlines which aim was to increase the specificity of evaluation of candidates for studies of the 2nd degree at the faculty of Dietetics and Food Science (D&FS). From the point of view of selecting the best candidates, criteria that are more beneficial are such that are characterised by considerable specificity, so as to avoid a situation in which there are people insufficiently prepared for studies among the students beginning their university career.

### Aim of study:

The aim of the presented analysis was a comparative evaluation of quality of two consecutive entry exam editions for studies of the 2nd degree at the faculty of D&FS at MUW, which was carried out in 2012 and 2103.

### Materials and Methods:

Entry exam for studies of the 2nd degree were analysed at the faculty of D&F carried out in 2012 (n = 83) and 2013 (n = 104), which comprised of the score from the test and GPA value from the period of studies of the 1st degree. In order to perform the discrimination analysis, in each of the MCQs exams three groups of questions were identified, according to the topic: human nutritional sciences, clinical outline of diseases, and clinical nutrition.

Evaluation of the test exam quality was performed establishing item easiness and discriminating power of the individual questions and distractors and to determine reliability of the test, Cronbach's alpha was determined. To evaluate dependency between score achieved in a certain area of knowledge and a total score in the exam, r-Pearson linear correlation coefficient was used. To evaluate discrimination power of the identified thematic groups and GPA value, Receiver Operating Characteristics (ROC) curve analysis was performed for two categories accepted/not accepted.

### Results:

The analysis of two editions of the exam shows quite significant differences in reliability of a test questions set ( $\alpha$  coefficient in 2012 and 2013 was 0.682 and 0.803, respectively). r-Pearson correlation coefficient for questions in clinical nutrition was greatly improved in 2013 when this set was questions in MCQs was increased (r-Pearson was 0.492 and 0.908, respectively). Moreover, it was observed that a change in question proportion in 2013 resulted in a more even spread of discrimination power on all three thematic areas. When comparing recruitment in 2012 and 2013 it was observed that GPA values are a weak selective factor, which is particularly visible in the edition of 2013 (AUC in 2012 and 2013: 0.6221 and 0.5109, respectively).

### Conclusions:

It is recommended that further evaluation of consecutive editions of an exam for the 2nd degree of studies be performed at the faculty of D&FS, by fine-tuning details of test content outlines. It is also recommended to evaluate selective efficiency of GPA further. Constant monitoring of students'

progress should be performed as well as graduates' careers of those who completed their education at D&FS at MUW.

Keywords: admission criteria, admission test, admission policy, grade point average, dietetics and food science.

## 1 INTRODUCTION

Following the provisions of the Bologna Declaration, the system of education in the European Union, including Poland, comprises two main cycles: the first-degree (bachelor) and the second-degree (master). Successful completion of a 3-year Bachelor programme allows for continuation of education on a 2-year Master programme [1, 2]. In Poland, beginning in 2005, candidates are admitted to a Bachelor's degree programme on the basis of the results of their matura exam (equivalent to A-levels in UK, Baccalauréat in France or Abitur in Germany). Matura results entitle secondary school graduates to apply to any of the Polish university-level schools, including 11 medical universities. Many countries have similar admission systems that are based on the assessment of candidates by grade point averages (GPAs) or General Certificate of Education Advanced Level (e.g. A-level), and aptitude tests [3]. In the academic year 2000/01, the Medical University of Warsaw (MUW) introduced a Bachelor's degree programme for the first time in its history, and a Master's degree programme was introduced in 2003/04. Only the so-called traditional faculties of medicine, dentistry, and pharmacy still offer integrated Master's programmes (one-cycle system). As an effect of the Bologna Declaration, a two-cycle education system was introduced at the remaining faculties comprising the field of health sciences (nursing, midwifery, public health, Dietetics and Food Science) [4].

The implementation of the provisions of the Bologna Declaration forced the development of a new system of higher education and introduction of an admission policy, ensuring a fair system of selection of the best candidates for a Master's degree programme. Unlike in the case of the admission process for a Bachelor's degree programme that is based on matura results, each university-level school had to develop certain selection methods for candidates for a Master programme. On the one hand, admission criteria are supposed to ensure equal conditions and fair assessment of each candidate; on the other hand, the criteria need to be reliable and valid in evaluating the features that are important for effective studying on a Master's degree programme [5]. Due to the fact that a Master programme is offered to a smaller number of students than a much more common Bachelor programme, searching for an effective selection tool to find the best candidates who, more likely than not, shall achieve professional success seems to be justified [6]. This is of particular importance in educating health care professionals since providing health security, which is one of the core functions of a state, requires well-prepared medical staff [7], as well as medical support personnel (health occupations). Dietician is one of health science professions. Graduates in Dietetics and Food Science from university-level schools are trained, among others, to work in health care centres (e.g. hospitals, outpatient clinics and units) and catering establishments (e.g. school canteens) as well as in centres that prepare and deliver food products (e.g. catering for nursing homes). The scope of professional activities mentioned above is typical of Bachelor degree graduates. Master's degree graduates obtain a considerably larger range of competence and they are supposed to be prepared to conduct research on assessment of dietary patterns and nutritional status as well as to monitor food quality, food storage conditions and production of meals. A Master's degree programme also prepares students to work in scientific and research institutes and research and development centres as well as institutions specialising in nutrition counselling and dissemination of knowledge regarding human nutrition. It is clearly seen that Master's students face significantly greater demands compared to Bachelor's students. This has to be considered when developing the selection tools. Therefore, between 2012 and 2013, the MUW made some changes in the test content outlines, aiming to increase the specificity of the assessment of candidates for a Master's degree programme in Dietetics and Food Science. Criteria of significant specificity are most valuable for the selection of best candidates so that persons with insufficient level of initial competences are not admitted to the course.

Considering the above, it is justified to pose a question whether the selection criteria for candidates for a Master's degree programme in Dietetics and Food Science at the MUW meet the assumptions of reliable and valid assessment. The aim of this analysis was to compare and assess two consecutive exam editions with respect to discriminatory capacity of selected thematic fields representing issues related to human nutrition. Evaluation comprised also the two remaining selection criteria used at the MUW, i.e. GPA and total point score on the admission test.

## 2 MATERIAL AND METHODS

The analysis involved the results of the entrance exams for a Master's degree programme in Dietetics and Food Science conducted in 2012 ( $n = 83$ ) and 2013 ( $n = 104$ ), which comprised the test score and GPA in Bachelor programme. The multiple-choice question (MCQ) exam was performed in the form of a test composed of 50 questions in the "best answer from a list of possible answers" format. Three varieties of a test that differed from each other only in the order of questions were prepared for each edition of the exam. For the purposes of factor analysis, three groups of questions were distinguished in each MCQ: *human nutritional sciences*, *clinical outline of diseases*, and *clinical nutrition*. See Table 1 for a list of the proportions of questions from particular thematic fields.

Table 1. Proportion of questions from particular thematic fields in accordance with test content outlines

MCQ exam	Human nutritional sciences	Clinical outline of diseases	Clinical nutrition
2012	20 (40,0%)	13 (26,0%)	10 (20,0%)
2013	14 (28,0%)	10 (20,0%)	21 (42,0%)

Scores on MCQ exams were pre-processed using TESTY version 7 ("Testy komputerowe", Copyright © 1994-2014 by Sławomir Zalewski), which allowed for calculation of quality parameters of the test questions. Information was read in the database, processed in Microsoft Excel 2010 (Microsoft Corporation) and exported to Statistica version 10 (StatSoft, Inc.) for further analysis. The software was used in compliance with the MUW license.

Normal distribution parameters of particular exam results were assessed using the Shapiro-Wilk test and data were screened for outliers using Grubbs test. One-way analysis of variance (ANOVA) was used to assess the equivalence of particular editions of the test exam. The quality of the exam was assessed by establishing the item easiness and discriminating power of particular questions and distractors, and the reliability of the exam was assessed by estimating Cronbach's  $\alpha$  coefficient (the Kuder-Richardson Formula 20 for a test comprising dichotomous variables) [8]. An analysis of parallel forms of the tests used in consecutive years was conducted by determining the Guttman split-half correlation [9]. Pearson's  $r$  linear correlation coefficient was used to assess the correlation between the score obtained for a certain thematic field and the total point score.

A non-parametric median comparison Mann-Whitney U test was used for comparative analyses of independent samples. ROC (Receiver Operating Characteristics) curve analysis for two categories (admitted/not admitted) was performed to estimate the discriminatory capacity of the thematic fields and GPA and a modified  $U$ -statistic as proposed by Hanley & Hajian-Tilaki was conducted to compare AUC (Area Under the Curve) [10, 11].

For all analyses, the relevance level assumed *a priori* was  $\alpha = 0.05$ .

## 3 RESULTS

The analysis of the parameters of distribution of the frequency of results of test exams demonstrated greater symmetry in the scores obtained in 2013, evidence of which can be found both in the value of skewness (slightly negative skew) and in a normal distribution of the data confirmed by the Shapiro-Wilk test. A comparison of the measures of central tendency showed that in the 2013 exam the most commonly occurring value of scores shifted towards the mean values of the point range, compared to the 2012 exam. See Table 2 for a detailed characteristics of both exams.

A comparison of scores obtained in 2012 and 2013 demonstrated that the 2013 scores had a wider scope of variability, which was an evidence of a greater differentiation of MCQ exam results among this group of candidates (Levene's test for homogeneity of variances,  $P = 0.04$ ). A median comparison Mann-Whitney U test also compared this difference ( $P = 0.0005$ ).

Due to the fact that there were 3 varieties of each 50-question exam, an analysis of equivalence was performed. No statistically significant differences were found between exam varieties in the two years that underwent analysis (ANOVA,  $P > 0.05$ ). Moreover, no outliers were found among the exam results (Grubbs test,  $P > 0.05$ ). This confirmed the homogeneity of results and equality of test conditions regardless of a test variety.

Table 2. List of parameters of quantitative assessment of tests conducted between 2012-2013

	2012	2013
Number of questions	50	50
Number of options	4	4
Number of varieties	3	3
Equivalent varieties*	$P > 0.05$	$P > 0.05$
N	83	104
Normal distribution**	$P < 0.05$	$P > 0.05$
Data outliers***	$P > 0.05$	$P > 0.05$
Mean (95% CI)	32.8 (31.6 – 34.0)	29.3 (28.0 – 30.6)
SEM	0.613	0.667
Median	34.0	29.5
Range of scores	16.0 – 42.0	13.0 – 42.0
Q <sub>1</sub>	30.0	25.0
Q <sub>3</sub>	37.0	35.0
SD (95% CI)	5.59 (4.85 – 6.59)	6.80 (5.98 – 7.87)
CV	17.05%	23.18%
Kurtosis	0.19	-0.67
Skewness	-0.68	-0.29

\* ANOVA test (for  $P < 0.05$  varieties are not equivalent)

\*\* Shapiro-Wilk test (for  $P < 0.05$  distribution is not normal)

\*\*\* Grubbs test for outliers (for  $P < 0.05$  presence of outliers)

CI – confidence interval; SEM – standard error of the mean; Q<sub>1</sub> – first quartile; Q<sub>3</sub> – third quartile; SD - standard deviation; CV - coefficient of variation

Reliability coefficient (Cronbach's  $\alpha$  coefficient) is the most frequently used method to assess internal consistency of measurement for at least two factors. The analysis of both exam editions demonstrated quite significant differences in reliability of test questions ( $\alpha$ -coefficient of 0.682 and 0.803, in 2012 and 2013 respectively). A detailed question evaluation allowed for the selection of questions with a negative correlation, i.e. significantly decreasing the reliability of the exam. With these questions excluded from the exam set, the overall reliability of tests increased ( $\alpha$ -coefficient amounted to 0.746 and 0.826 in 2012 and 2013, respectively). Apart from differences in the level of reliability, the year 2013 witnessed an increase in the general difficulty of the exam compared to 2012 (average item easiness amounted to 0.655 and 0.587 in 2012 and 2013, respectively). There was also an increase in the test capacity of differentiating candidates in 2013 (mean differentiation power of questions amounted to 0.252 and 0.308 in 2012 and 2013, respectively). See Table 3 for a list of the results of analysis of reliability, easiness, and differentiating capacity.

Table 3. Results of analysis of reliability and qualitative parameters for exam tests conducted between 2012-2013

	2012	2013
Cronbach's $\alpha$	0.682	0.803
Number of questions needed to achieve $\alpha = 0.85^*$	132	70
Number of questions with negative correlation	6	5
Cronbach's $\alpha$ after optimization	0.746	0.826
Index of test difficulty	0.655	0.587
Average discriminatory power	0.252	0.308

\* calculated using the Spearman–Brown prediction formula

Guttman split-half correlation was determined to assess the reliability of the two parallel forms of tests used in subsequent exam editions. In order to select test halves, a random sampling of 25 questions was performed using stratification that involved dividing the pool of questions into three thematic areas, and split-half variances were subsequently calculated (9.9113 and 17.1273 in 2012 and 2013, respectively). With the total variance of 42.2855, the coefficient of equivalence of test versions measured by the Guttman's technique amounted to 0.721.

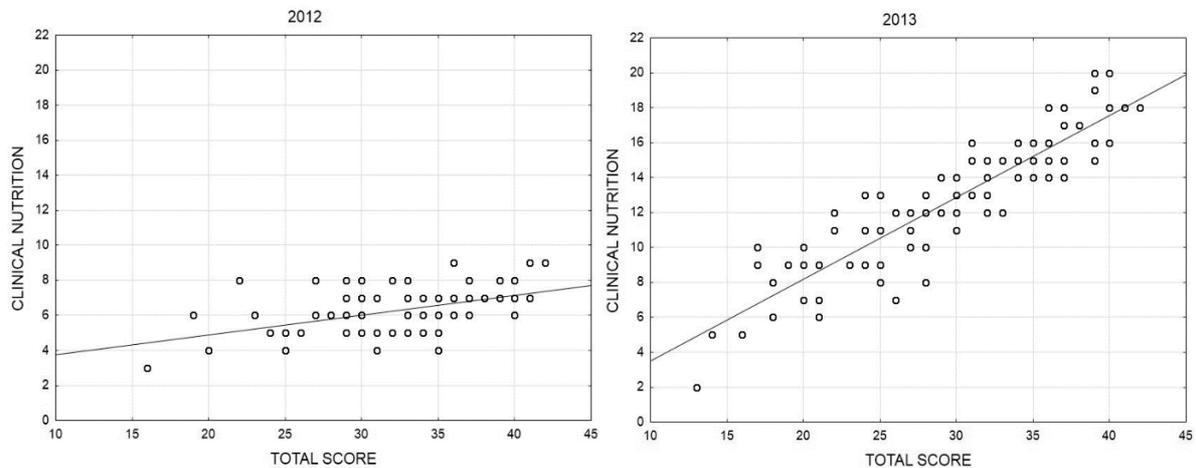


Figure 1. Effects of change in proportion of questions regarding *clinical nutrition* on correlation between results in this thematic field and total exam score. Pearson's *r* correlation coefficient amounted to 0.492 in 2012 and 0.908 in 2013, respectively

An adjustment to the test content outlines, performed between 2012 and 2013 and consisting in a change in the proportion of questions from particular thematic fields (see Table 1), changed the strength of correlation between the results obtained for a particular group of questions and the total score obtained on the MCQ exam. Table 4 shows that Pearson's *r* linear correlation coefficient for the questions regarding *clinical nutrition* significantly improved in 2013 when the proportion of this pool of questions in MCQ exam had increased (correlation coefficient of 0.492 and 0.908 in 2012 and 2013, respectively). Figure 1 shows the change in the strength of correlation between the thematic groups and total point score in 2012 and 2013.

Table 4. Correlations between results obtained for particular thematic fields and total point score on entrance exam

		Human nutritional sciences	Clinical outline of diseases	Clinical nutrition
<b>2012</b>	Average scores	13.3	9.9	6.3
	SD	2.37	2.70	1.29
	R	0.841	0.888	0.492
	$r^2$	0.707	0.789	0.242
	P-value*	0.0000001	0.0000001	0.000002
<b>2013</b>	Average scores	8.3	6.1	12.6
	SD	2.23	1.78	3.52
	R	0.806	0.724	0.908
	$r^2$	0.650	0.524	0.824
	P-value*	0.0000001	0.0000001	0.0000001

SD - standard deviation; Pearson's *r* linear correlation coefficient;  $r^2$  – coefficient of determination  
\* testing using Student's t-distribution

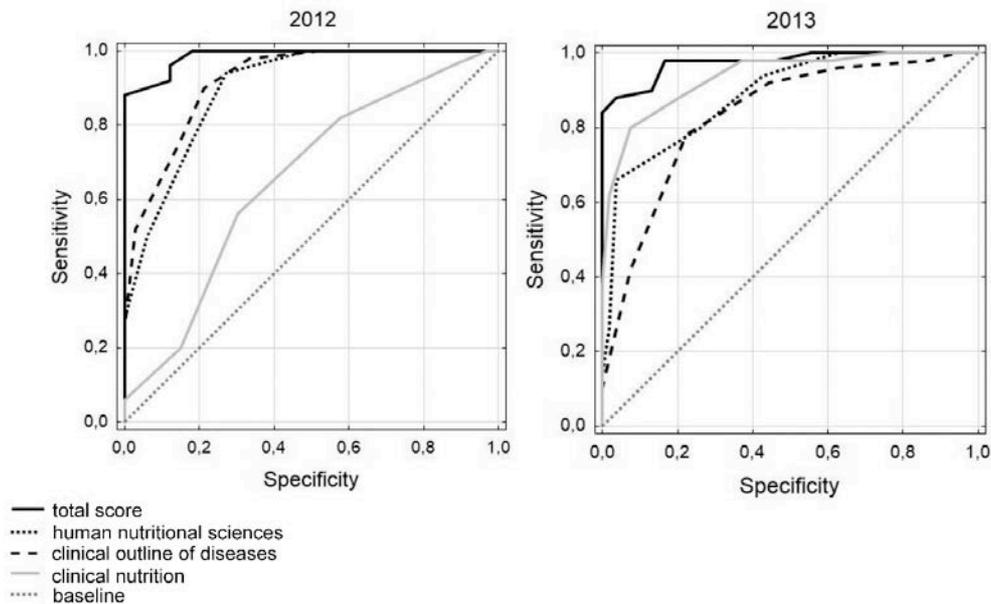


Figure 2. Assessment of discriminatory capacity of selection criteria, including thematic areas distinguished for test exam, established based on ROC curves

ROC curves were compared to analyse the effects of adjustments to the test content outlines on discriminatory capacity of the three thematic areas represented in the exam test. Figure 2 demonstrates that the change in the proportion of questions in 2013 contributed to a more even distribution of discriminatory power to all three thematic areas. An insignificant reduction in discriminatory capacity of questions regarding *human nutritional sciences* and *clinical outline of diseases* produced a high discriminatory level for the third thematic field, i.e. *clinical nutrition* (AUC of 0.6545 and 0.9359 in 2012 and 2013, respectively).

ROC curves analysis was also used to assess the selection capacity of the second criteria, i.e. GPA in Bachelor programme. In both editions of the entrance exam, GPA was of the same importance for the total score. The comparison of 2012 and 2013 exams showed that GPA was a poor selection criterion, which was particularly visible in the 2013 edition (AUC of 0.6221 and 0.5109 in 2012 and 2013, respectively).

#### 4 DISCUSSION

The issue of proper adjustment of admission policy to the needs of programmes in medicine-related majors has been analysed for decades [5, 12, 13]. It would seem that it is sufficient to use verified and validated selection tools, which is not supposed to cause major analytical difficulties. Meanwhile, experience has shown that this is a significant problem that has not been solved yet. A statement from one of the publications by Judy Searle & Jane McHarg can be quoted here: "*Just pick the right students and the rest is easy!*" [14]. The question "*Who are the most suitable candidates for medical university?*" is still valid [15].

These issues have been analysed and evaluated by numerous researchers, in particular with reference to the selection of proper candidates for a programme in medicine and less with reference to other medical professions such as dentistry [16] and pharmacy [17, 18]. There is a number of publications on good practices in selection of candidates for studying health professions such as nursing and midwifery [19]. There are no reliable and long-term analyses on admission strategies in the case of selection of proper candidates for a programme in Dietetics and Food Science. A development of sufficient selection tools for this major requires reference to certain solutions and good practices developed for other health occupations.

From among numerous and diverse criteria, in their admission procedures, most medical schools use GPA as well as standardised aptitude tests, mini-interviews and written assignments such as essays [20]. Several analyses confirmed that GPA calculated for the period preceding the start of a MA degree programme as well as grades obtained during a BA degree programme are good predictive factors that allow for an accurate assessment of the predispositions of prospective students of

medicine or nursing [21-23]. The analysis demonstrated that GPA does not have a profound influence on discriminatory capacity of human nutrition candidates. ROC curves analysis showed that the selective capacity of GPA was low, and it was statistically insignificant in 2012-2013, falling far behind the scores on MCQ exams. Many authors have pointed to the problem of effectiveness of the selection based on GPA due to the heterogeneity of grading systems in various educational institutions; therefore, the reliability of selection of candidates based on GPA needs to be regarded as poor [24, 25]. If, despite the unfavourable results described above, GPA is preserved as a selection criterion for a Master's degree programme, its importance for the total score should be estimated carefully. Attaching too much importance to GPA may contribute to favouring candidates with BA degrees who graduated from universities with insufficient quality of their assessment systems. In compliance with a critical analysis by McNelis et al. regarding the admission policy for a MA programme in Nursing, it can be assumed that the appropriate significance level for each criterion is supposed to be as follows: GPA (30%); nursing GPA (30%); interview (30%); and service experience (10%) [24]. Observation needs to be continued in the subsequent years to assess the predictive capacity of GPA to predict future careers of Dietetics and Food Science students and graduates.

The issue of efficiency of different selection methods based on various admission tests is most widely described in the available literature. The following well-known medical examinations can serve here as an example: Medical Colleges Admission Test (MCAT) [26], the UK Clinical Aptitude Test (UKCAT) [27], the Graduate Australian Medical School Admissions Test (GAMSAT) [28], and the Irish Health Professions Admission Test (HPAT) [29]. Evaluation analyses related to other admission tests are less available, for example Nursing Entrance Test (NET) [30], and Health Education Systems, Inc Admission Assessment Exam (HESI A<sup>2</sup>) [31]. Chen & Voyles emphasised that an entrance test exam is a good selection criterion allowing for choosing the best candidates for the course. Positive results of correlation analyses and good predictive value of this criterion are important reasons supporting the thesis above since MCQ exam allows, with high probability, for prediction of future career of students and graduates [25, 32]. It needs to be emphasised that the above mentioned admission tests basically refer to the selection of candidates applying to a medical university straight after graduation from a secondary school (BA degree programme or integrated programmes). These well-studied tools cannot necessarily be directly applied to the admission procedure for a Master programme.

Reliability (understood as the repeatability of results obtained under certain conditions) is supposed to be an essential element of a good selection procedure [33]. Insufficient reliability of qualification procedures contributes to a low degree of confidence since particular results differ markedly in similar conditions. The present analysis of reliability conducted with the use of Cronbach's  $\alpha$  coefficient demonstrated that the change in test content outlines in 2013 contributed to a significant increase in reliability ( $\alpha$  coefficient > 0.8). These positive results may be associated with an increase in domain coherence of the question set. A further improvement of reliability parameters of the MCQ exam may be achieved by extending the measurement scale. Calculations with Spearman-Brown prophecy formula showed that an increase in reliability of measurement may be achieved by adding more questions to the pool. Obviously, seeking exclusively high values of  $\alpha$  coefficient does not guarantee exam reliability, since a high value of  $\alpha$  coefficient only minimizes the impact of random errors on the results and does not guarantee the lack of systematic errors that might be even serious [34].

Annual performance of the entrance exam in accordance with the modified admission procedures has to account for the fact that the subsequent exams are supposed to keep high reliability of assessment. Guttman split-half correlation was used to compare MCQ exams with reference to meeting the criterion of parallel forms [9, 35]. Guttman correlation coefficient for random test halves with domain stratification of questions showed a medium correlation of both variants, which proved that the qualitative parameters of the 2013 test differed significantly compared to those of the 2012 edition. Thus, the reliability of the exams differed considerably, which can be attributed to the increase in domain coherence of the 2013 test compared to the 2012 edition. This was also confirmed by a much more balanced correlation between the scores obtained for each thematic group and the total point score in 2013. These observations proved that the adjustments introduced to the test content outlines promoted the increase in reliability of the measurement.

The comparison of admission tests demonstrated an increase in general difficulty of the MCQ exam as well as the mean differentiation power of questions in 2013 compared to 2012. The scope of variability of results and the symmetry in test results were greater. All those factors proved that the quality of the exam improved; the exam itself has become more selective, facilitating the process of discrimination of candidates. Thus the thesis that a proper selection of exam questions in terms of quality is an important factor that needs to be considered in development of the MCQ exam was strongly confirmed

[36]. A large number of easy and approving questions that require only confirmation of correctness of given information (e.g. the use of "all of the above" (AOTA) and "none of the above" (NOTA) as options) [37] is intellectually less demanding and contributes to obtaining better scores [38]. Therefore, it is important to choose exam questions with a significant differentiation power. These questions are capable of distinguishing those candidates who obtain significantly different results for a particular parameter, which allows for a good selection [39]. Such tasks are of particular value since they facilitate the assessment of practical aspects of using knowledge, understanding complex issues, and familiarity with the mechanisms and the course of biological processes. In the literature, these tasks are referred to as "burst tasks" [39], since they cause the flatness of the distribution of results (curtosis  $< 0$  in the 2013 edition), which facilitates discrimination and establishment of the "passed/failed" cut-off point. The comparative analysis demonstrated that the 2013 exam had significantly better qualitative characteristics measured with item easiness and differentiation power of the exam questions. This positive experience may become the foundation of good practice in the development of subsequent editions of the exam. The development of a proper database of exam questions is a long-lasting process and database resources need to be successively updated due to the process of "ageing" of exam questions. This process consists in an increase of easiness and decrease of differentiation power in the subsequent editions comprising the database [40].

Apart from the issue of reliability ("how is it measured?"), the determination of validity of measurement ("what is measured?") is also crucial for a development of good assessment tools. Validity is here understood as the usefulness of a particular tool for assessment of a certain set of features and qualities of candidates [41]. The selection of objective criteria used in an admission process is necessary to conduct a reliable selection. If we want to have a good candidate selection procedure, we follow the principle of choosing those who have possessed a minimum of knowledge and skills necessary for a certain profession and represent certain predispositions for this particular profession. Therefore, if we use the MCQ exam score as the main selection criterion, this tool should accurately assess those initial competences that are essential for taking a Master programme in Dietetics and Food Science. The ROC curves analysis showed that the change in test content outlines in particular thematic fields had a positive influence on the discriminatory capacity. A comparable selection capacity of each thematic group was obtained due to the modification of proportion of particular questions in certain thematic areas. Incorrect proportion of particular topics and excessive representativeness of tasks referring to one thematic field strongly influenced the validity of assessment conducted with the MCQ exam. Internal validity (known also as content validity) of a particular tool may only be achieved by developing a strict test content outlines [42]. Implementation of the outlines in the subsequent editions of admission tests shall allow for validation of assumptions and their potential improvement in the next years.

The present analysis involved an assessment of only a few principal factors that could determine an admission procedure of high quality. A number of factors should be considered in the principles of selection of the best candidates. It is advised to assess both the "cognitive" and "non-cognitive" criteria [43], as recommended by the following institutions: The American Psychological Association, American Educational Research Association, and National Council on Measurement in Education's Standards for Educational and Psychological Testing [41]. A multifactorial approach to assessing applicants has been espoused by many researchers, citing the need to assess both 'cognitive' and 'non-cognitive' criteria to better choose candidates that will not only succeed academically and professionally, but also to minimize attrition by choosing well-suited and motivated candidates [43-45]. However, the use of a complex measurement tool impedes an independent assessment of selection efficacy of the chosen components. If we want to have a comprehensive, evidence-based admission policy, we need to have an opportunity to validate the tools used for the selection procedure. Therefore, the question asked by Wilson 15 years ago: "*How can the process we had always used be improved?*" is still valid [46].

## 5 CONCLUSIONS

The present analyses and critical assessment of the results allow us to propose a number of recommendations for the admission strategy for a Master's degree programme in Human Nutrition at the MUW:

1. Maintenance of the present two selection criteria used in the admission procedure for a Master's programme in Dietetics and Food Science;
2. Introduction of a "non-cognitive" criterion into the admission system;

3. Development of a database of exam questions through systematic collection of exam questions and assessment of their quality as well as performance of standardization of exam tasks from selected fields of knowledge and skills subdivided into certain domains;
4. Continuation of monitoring of careers of Dietetics and Food Science students and graduates in order to gather information necessary to conduct a predictive analysis of the selection methods used in this major.

## REFERENCES

- [1] Cumming, A. (2010). The Bologna process, medical education and integrated learning. *Medical Teacher* 32(4), pp. 316-318.
- [2] Fernandez, F.A. (2010). Adaptation of the new medical degree to the European Space for Higher Education. What has been the contribution of Bologna? *Revista Clinica Espanola* 210(9), pp. 462-467.
- [3] McManus, I.C., et al. (2005). Unhappiness and dissatisfaction in doctors cannot be predicted by selectors from medical school application forms: a prospective, longitudinal study. *BMC Medical Education* 5, pp. 38.
- [4] Kraśniewski, A., *Proces Boloński - to już 10 lat. 2009*, Warsaw: Fundacja Rozwoju Systemu Edukacji.
- [5] Roberts, C. and D. Prideaux. (2010). Selection for medical schools: re-imaging as an international discourse. *Medical Education* 44(11), pp. 1054-6.
- [6] Marz, R., et al. (2013). Tuning research competences for Bologna three cycles in medicine: report of a MEDINE2 European consensus survey. *Perspectives on Medical Education* 2(4), pp. 181-95.
- [7] Weissert, C.S. and S.L. Silberman. (1998). Sending a policy signal: State legislatures, medical schools, and primary care mandates. *Journal of Health Politics Policy and Law* 23(5), pp. 743-770.
- [8] Feldt, L.S. (1969). A test of hypothesis that Cronbachs alpha or Kuder-Richardson coefficient 20 is same for 2 tests. *Psychometrika* 34(3), pp. 363.
- [9] Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika* 10(4), pp. 255-82.
- [10] Wanvarie, S. and B. Sathapatayavongs. (2007). Logistic regression analysis to predict Medical Licensing Examination of Thailand (MLET) Step1 success or failure. *Annals of the Academy of Medicine, Singapore* 36(9), pp. 770-3.
- [11] Hanley, J.A. and K.O. Hajian-Tilaki. (1997). Sampling variability of nonparametric estimates of the areas under receiver operating characteristic curves: an update. *Academic Radiology* 4(1), pp. 49-58.
- [12] Vaughn, K.W. (1946). The graduate record examination as an aid in the selection of medical students. *Journal of the Association of American Medical Colleges* 21, pp. 129-46.
- [13] Benbassat, J. and R. Baumal. (2007). Uncertainties in the selection of applicants for medical school. *Advances in Health Sciences Education: Theory and Practice* 12(4), pp. 509-21.
- [14] Searle, J. and J. McHarg. (2003). Selection for medical school: just pick the right students and the rest is easy! *Medical Education* 37(5), pp. 458-463.
- [15] Cowley, C. (2006). Polemic: five proposals for a medical school admission policy. *Journal of Medical Ethics* 32(8), pp. 491-494.
- [16] Ranney, R.R., M.B. Wilson, and R.B. Bennett. (2005). Evaluation of applicants to predoctoral dental education programs: review of the literature. *Journal of Dental Education* 69(10), pp. 1095-106.
- [17] Latif, D.A. (2005). Including the assessment of nontraditional factors in pharmacy school admissions. *Annals of Pharmacotherapy* 39(4), pp. 721-6.

- [18] Panczyk, M., H. Rebandel, and J. Gotlib (2015). Assessment of predictive value of admission criteria of candidates for pharmaceutical studies - an empirical investigation. *Indian Journal of Pharmaceutical Education and Research* 49(2), pp. 112-120.
- [19] Sabin, M., R. Taylor, and C. Tilley. (2012). Untangling a complex issue: An overview of initiatives to support nursing and midwifery student recruitment, selection and retention in Scottish Universities. *Nurse Education Today* 32(4), pp. 469-474.
- [20] Monroe, A., et al. (2013). An overview of the medical school admission process and use of applicant data in decision making: what has changed since the 1980s? *Academic Medicine* 88(5), pp. 672-81.
- [21] Ali, P.A. and P.B. Naylor. (2010). Association between academic and non-academic variables and academic success of diploma nursing students in Pakistan. *Nurse Education Today* 30(2), pp. 157-62.
- [22] Shulruf, B., et al. (2011). Rethinking the admission criteria to nursing school. *Nurse Education Today* 31(8), pp. 727-732.
- [23] Salvatori, P. (2001). Reliability and validity of admissions tools used to select students for the health professions. *Advances in Health Sciences Education: Theory and Practice* 6(2), pp. 159-75.
- [24] McNelis, A.M., et al. (2010). Revision and evaluation of the Indiana University School of Nursing baccalaureate admission process. *Journal of Professional Nursing* 26(3), pp. 188-95.
- [25] Chen, S. and D. Voyles. (2013). HESI Admission Assessment Scores: Predicting Student Success. *Journal of Professional Nursing* 29(2 SUPPL1), pp. S32-S37.
- [26] Albanese, M.A., et al. (2003). Assessing personal qualities in medical school admissions. *Academic Medicine* 78(3), pp. 313-21.
- [27] Lynch, B., et al. (2009). Does the UKCAT predict Year 1 performance in medical school? *Medical Education* 43(12), pp. 1203-9.
- [28] Coates, H. (2008). Establishing the criterion validity of the Graduate Medical School Admissions Test (GAMSAT). *Medical Education* 42(10), pp. 999-1006.
- [29] Kelly, M.E., et al. (2013). To what extent does the Health Professions Admission Test-Ireland predict performance in early undergraduate tests of communication and clinical skills? An observational cohort study. *BMC Medical Education* 13, pp. 68.
- [30] Ellis, S.O. (2006). Nurse entrance test scores: a predictor of success. *Nurse Educator* 31(6), pp. 259-63.
- [31] Murray, K.T., C.S. Merriman, and C. Adamson. (2008). Use of the HESI Admission Assessment to predict student success. *Computers, Informatics, Nursing* 26(3), pp. 167-72.
- [32] Lancia, L., et al. (2013). Academic success or failure in nursing students: Results of a retrospective observational study. *Nurse Education Today* 33(12), pp. 1501-1505.
- [33] Babbie, E. (2013). *The practice of social research*. 13th ed. Belmont: Cengage Learning.
- [34] Guilford, J.P. (1954). *Psychometric methods*. 2nd ed. New York: McGraw-Hill.
- [35] Green, S.B. (2003). A coefficient alpha for test-retest data. *Psychological Methods* 8(1), pp. 88-101.
- [36] Tarrant, M. and J. Ware. (2008). Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Medical Education* 42(2), pp. 198-206.
- [37] Harasym, P.H., et al. (1998). Cuing effect of "all of the above" on the reliability and validity of multiple-choice test items. *Evaluation & The Health Professions* 21(1), pp. 120-33.
- [38] Boland, R.J., N.A. Lester, and E. Williams. (2010). Writing Multiple-Choice Questions. *Academic Psychiatry* 34(4), pp. 310-316.

- [39] Downing, S.M. (2005). The effects of violating standard item writing principles on tests and students: the consequences of using flawed test items on achievement examinations in medical education. *Advances in Health Sciences Education: Theory and Practice* 10(2), pp. 133-43.
- [40] Case, S.M. and D.B. Swanson. (2002). *Constructing Written Test Questions For the Basic and Clinical Sciences*. 3rd ed. National Board of Medical Examiners.
- [41] Goodwin, L.D. (2002). Changing conceptions of measurement validity: an update on the new standards. *The Journal of Nursing Education* 41(3), pp. 100-6.
- [42] Ebel, R.L. (1961). Must all tests be valid? *American Psychologist* 16(10), pp. 640-647.
- [43] Wolkowitz, A.A. (2011). Multiple attempts on a nursing admissions examination: Effects on the total score. *Journal of Nursing Education* 50(9), pp. 493-501.
- [44] Kulatunga-Moruzi, C. and G.R. Norman. (2002). Validity of admissions measures in predicting performance outcomes: the contribution of cognitive and non-cognitive dimensions. *Teaching and Learning in Medicine* 14(1), pp. 34-42.
- [45] Timer, J.E. and M.I. Clauson. (2011). The use of selective admissions tools to predict students' success in an advanced standing baccalaureate nursing program. *Nurse Education Today* 31(6), pp. 601-606.
- [46] Wilson, T. (1999). A student selection method and predictors of success in a graduate nursing program. *Journal of Nursing Education* 38(4), pp. 183-7.