journal homepage: http://kont.zsf.jcu.cz
DOI: 10.32725/kont.2022.018

KONTAKT / Journal of nursing and social sciences related to health and illness

Original research article

# Michigan diabetes knowledge test: validation of modified version among Czech students of nursing

Pavla Kudlová [1] *, Ilona Kočvarová [2], Petr Soukup [3]

[1] *Tomas Bata University in Zlín, Faculty of Humanities, Institute of Health Sciences, Zlín, Czech Republic*
[2] *Tomas Bata University in Zlín, Faculty of Humanities, Research Center, Zlín, Czech Republic*
[3] *Charles University, Faculty of Social Sciences, Department of Sociology, Prague, Czech Republic*

**Abstract**
Michigan Diabetes Knowledge Tool (MDKT) was developed to test the level of knowledge in managing Diabetes Mellitus. It is widely used as an original (for patients) and modified tool (for nurses or nursing students) around the world.
*Objectives:* The aim is psychometric testing of the modified MDKT.
*Methods:* This quantitative study uses descriptive statistics, Classical Test Theory (CTT), Item Response Theory (ITT), and reliability estimation within the convenient sample of students of nursing ($N$ = 133). Data collection took place in 2018–2019 at a selected Czech University.
*Results:* Using the MDKT only enables us to differentiate within a narrow spectrum of students whose knowledge is sub-average to average. The analysis has suggested the possibility of reducing the test while maintaining its psychometric parameters.
*Conclusions:* The test is reliable, however too simple for our convenient sample of third-year nursing students, but this result cannot be generalised. Future investigations may lead to possible modifications, for example creating a more difficult version of the test and its re-validation.

**Keywords:** Classical Test Theory; Item Response Theory; Michigan Diabetes Knowledge Test (MDKT); Nursing students; Reliability; Validity

## Introduction

The development of our current society has shown an increasing prevalence of diabetes mellitus (DM). Thus it is necessary to pay greater attention to this problem at the scientific level. Education is one of the ways to do this, as it can be focused (among others) on nursing professionals (World Health Organization, 2020). Knowledge of diabetology needs to be tested, albeit within the course of study or the course of nursing practice. To this end, it is necessary to have a valid and reliable tool, usually a didactic test. In this study we use the terms test or tool, usually reflecting the Michigan Diabetes Knowledge Test. We consider both terms to be acceptable in this context.

The preparation and validation of tests is one of the classic disciplines, not only in teacher education, but also in all the other fields where amassed knowledge, skills, competences, or other constructs connected to human activity must be validated. It is the same in nonmedical health care subjects, where reliable knowledge-monitoring tools are necessary, typically on entering, during, and at the end of a study program, but also continuously during practice.

Within our review, carried out for the period of 1983–2020 in 5 scientific electronic databases, 2 most frequently used standardized tools for the assessment of nurses' knowledge of DM were identified. The tools had been repeatedly used to assess the knowledge of nurses working in various clinical conditions/areas.

The first tool – the Diabetes Knowledge Test (DKT) by Scheiderich et al. (1983) contained 33 items. In 1989 it was extended by Drass et al. to 45 items. This 45-item version has been reviewed several times in compliance with the contemporary guidelines in diabetology. These modifications began to be referred to as the Modified Diabetes Basic Knowledge Test (MDBKT). For example, the MDBKT was used to assess nurses' knowledge in the studies by Chan and Zang (2007) and Yacoub et al. (2014, 2015).

The second tool is the Michigan Diabetes Knowledge Test (MDKT) created in 1998 by Fitzgerald et al. The original MDKT was created by the Diabetes Research and Training Center at the University of Michigan. The original MDKT was focused on patients. The test, comprising 23 items, has been modified several times to comply with contemporary guidelines in diabetology. The number of items, however, has not been changed.

* **Corresponding author:** Pavla Kudlová, Tomas Bata University in Zlín, Faculty of Humanities, Institute of Health Sciences, Štefánikova 5670, 760 01 Zlín, Czech Republic; e-mail: kudlova@utb.cz | www.fhs.utb.cz
http://doi.org/10.32725/kont.2022.018
Submitted: 2021-07-31 • Accepted: 2022-05-30 • Prepublished online: 2022-06-09
KONTAKT 24/2: 114–122 • EISSN 1804-7122 • ISSN 1212-4117

The test is also known as the revDKT (Fitzgerald et al., 2016), but in our study we used only the MDKT abbreviation to make things simple. The MDKT includes 14 general knowledge multiple-choice items primarily meant for persons with Type 1 and Type 2 DM, and 9 multiple-choice items meant for persons treated with insulin (Fitzgerald et al., 1998). There is 1 correct answer to every question, and two or three distractors. For the needs of testing patients with DM it was translated to Spanish, Greek, Navajo, Norwegian, and Bahasa Malay. The test was also integrated into the automated medical record Kaiser Permanente Health Connect in California. The MDKT/revDKT is freely accessible, the only condition being to quote the original authors. In 2016, a 19-item "Simplified Diabetes Knowledge Scale (SDKS)" based on the MDKT by Fitzgerald et al. (1998) was developed by Collins et al. (2011). The scale is based on choosing the following answers – true, false, I don't know. The authors tested it on a group of patients with DM.

Although the test was originally intended for patients, it is also used to test medical staff. In 2016 it was used by researchers for testing nursing staff in Norway (Haugstvedt et al., 2016), and in 2020 researchers used it for testing nurses working in general practitioners' offices in the Czech Republic (Kudlová and Kočvarová, 2020). The test was also applied to nursing students from Japan and Australia (Ramjan et al., 2017), and to students of medicine (Beverly et al., 2019).

There are also other testing tools that have been used to test the level of nurses' knowledge independently (e.g., Hollis et al., 2014; Kudlová et al., 2017; Rhodes et al., 2019; Rubin et al., 2007), or in combination of the test itself and the standardized MDKT (Kudlová and Kočvarová, 2020).

For example, a 21-item self-constructed tool used in the research by Rubin et al. (2007) was based on the contemporary standards in diabetic foot care. It had a good inner consistency (Cronbach's $\alpha = 0.78$) and proved to be reliable in a repeated test (Pearson correlation coefficient of 0.71). The total mean success rate in the test on the given research set was 61%.

Hollis et al. (2014) used a 14-item multiple-choice knowledge test from the National Association of Diabetes Centers (NADC). The test's inner consistency was high (Cronbach's $\alpha = 0.94$).

Kudlová et al. published a study in 2017 in which they used a 40-item quasi-standardized test of own construction meant for the participants of a post-graduate course in diabetology. The final test (40 questions/semi-open tasks) was focused on assessing the level of acquiring knowledge in diabetology. The questions were created by experts taking part in the content of a certified course, and annually updated in terms of new findings in the field. The mean Cronbach's alpha reached 0.73 at the end of the course.

In 2019, Rhodes et al. published a study, the goal of which was to assess the efficacy of a quick e-learning module designated for professional development in diabetes management in school nurses in school. The pre-post-test included 15 multiple-choice items and 5 true-false items. The questions were derived from learning goals, information, and sources provided in the quick e-learning module. The test was statistically assessed, but its validity and reliability are not mentioned. The overall test results showed a statistically significant improvement between the pre- and post-tests.

In 2020, Kudlová and Kočvarová published a study focused on finding the amount of knowledge and the level of self-assessment of nurses working in primary outpatient care in the area of DM. The questionnaire contained a total of 68 items and was composed of four parts. Part 3 included a 23-item MDKT, complemented with 10 semi-open items created for the assessment of the certified course participants' knowledge, focused on nursing care and education in diabetology with a mean Cronbach's alpha of 0.76.

The MDKT is frequently used. There are multiple language mutations, and it is applied across various groups of respondents. Continuous attention needs to be paid to its modification and validation in various contexts to prevent its inappropriate application, and ideally maximize its usefulness.

The overall goal of this study is to validate a modified version of the Michigan Diabetes Knowledge Test (MDKT) as applied to third-year nursing students of a Czech university. The overall goal is divided into four sub-goals. We aim to (1) find out and assess the overall results of students in the test; (2) assess the psychometric properties of individual items (difficulty, sensitivity, and analysis of distractors) based on Classical Theory of Tests (CTT); (3) assess the psychometric properties of individual items (probability of correct answer, parameter estimates and item fit) based on Item Response Theory (IRT); (4) estimate reliability of the tool.

## Materials and methods

### *Sample and timing of the research*

The validation study is based on a convenience research sample (the participants were selected based on their availability and willingness to take part in the research). The sample included 133 students of the Bachelor program of Nursing. These were students of the third (graduate) year (70 full-time and 63 part-time students). All students were from one Czech university. Our research was conducted at the turn of 2018–2019. First, translation and modification of the tool was realized (2018). Second, selection of the research sample and data collection was realized, taking into account ethical aspects of the research (2019).

### *Research tool and its modification*

We applied the MDKT, namely its reviewed version (Fitzgerald et al., 2016). The test exists in several language versions, as mentioned above. Within our study, the test was professionally translated to Czech and partially modified.

The translation of the MDKT was carried out in several stages. The requirements were set by the authors in the following way: (1) independent translations were done by two native speakers from English to Czech; (2) independent reverse translations were done from Czech to English; (3) tabulation of all translations was done pointing out the differences; (4) discussion among translators and experts in the field of diabetology aimed at finding consensus; (5) final proofreading and creating the final version (Mandysová, 2019).

The modification was realized based on expert assessment of each item of the test, and also with respect to our plan to apply it in the Czech environment to students of nonmedical subjects (because the original tool is designed for patients). Our modification of the test lies in reformulating the items so that they are not applied directly to diabetics, but also other groups of respondents (in our case students of nursing). This modification concerns a total of 8 items (5, 11, 12, 16, 17, 18, 19, and 21).

The items were also modified in compliance with the current trends and recommendations (ADA, 2019; Bergenstal et al., 2010; CDS, 2020; Christiansen et al., 2018; Davies et al., 2018). Up to 92% persons with DM in the Czech Republic do not measure keto substances. It is preferred to monitor glycemia using a glucometer, or the so-called Continuous Glucose

Monitoring System (CGMS). Both these systems are easily available in the Czech Republic. According to these recommendations, Item 4 was significantly changed. Although this change is significant and alters the overall spirit of the item, it seemed to be necessary. Another modification concerns a better specification of the terms used in the test so that they were more unambiguous [2: Swiss cheese – 30% cheese Eidam; 3: low fat milk – low fat milk (2%); 8: a glass of diet drink – a glass of diet drink (e.g., Coke Light); 19: some juice – sweetened drink/juice].

A pilot study was carried out on the sample of 6 respondents (3 students of the General Nursing program and 3 students of the Registered Nurse adaptation program). They were asked to fill out the MDKT and to make possible comments on the content and formal aspect of the Czech version of the test. No comments were made that would require the reworking of the test items.

The test, in its applied form (after its modification), is part of the appendix (Suppl. 1). The items of the text are listed using their number in the test.

### *Analysis*

For the purposes of this analysis, the answers were recoded into binary variables (1 = correct, 0 = incorrect). The total score of each respondent was calculated and then recalculated in the form of success rate (percentage). There are no missing values in our data set. The statistical analysis is based on descriptive statistics, CTT (difficulty, sensitivity, and an analysis of distractors), IRT (probability of correct answer, parameter estimates and item fit), and reliability estimation. The analysis was realized using the SPSS version 25 for the purposes of descriptive statistics. Other parts of the analysis were realized using free software: R version 3.6.2 and its package ShinyItemAnalysis version 1.3.4. (SIA; Martinková and Drabinová, 2018). The full study protocol is available on request from the first author of the study.

We believe that a combination of these approaches enables us to evaluate the validity and reliability of the tool when applied to our research sample (without the possibility of generalization to the whole population).

## Results

Before moving to the substantive results, we present a more detailed description of our research sample. The respondents' age was 21–54 years, 26.4 on average (SD = 7.1). The students had dealt with the problem of DM in clinical subjects (71%), in nursing practice (70%), in the subject of Diabetology (52%), within basic subjects (50%), in nursing subjects (48%), or within self-study (30%) or scientific courses (12%). Most of the students (65%) stated that within their practice they had treated over 50 patients with DM. At the same time all of them (100%) measured glycemia using a glucometer, almost all (over 95%) administered medicines, applied insulin or took samples of blood or urine for the laboratory. 88% served food, 87% monitored physiological functions, 80% redressed wounds, and 63% educated patients. These details about the research sample are included here to point out its specificity. These were students at the end of their studies who were already active in practice and could be considered as fledgling medical staff. It is apparent that those were relatively knowledgeable and experienced respondents who were well-oriented in this field. This fact needs to be taken into account while interpreting further results of this study.

To follow the goal of this study and its four sub-goals, the results are further divided into four parts: (1) descriptive results, (2) psychometric properties based on CTT; (3) psychometric properties based on IRT; (4) reliability estimate.

### *Overall results of students in the test*

Table 1 shows the overall results of students in the test that are needed for the purposes of basic evaluation of the MDKT.

**Table 1. Descriptive results of students in the test**

| Test battery | *N* | Minimum | Maximum | Mean | St. deviation |
|---|---|---|---|---|---|
| Part 1: overall score | 133 | 2 | 14 | 10.35 | 3.10 |
| Part 2: overall score | 133 | 2 | 9 | 7.50 | 1.80 |
| Overall test score | 133 | 7 | 23 | 17.85 | 4.47 |
| Part 1: success rate (%) | 133 | 14% | 100% | 74% | 0.22 |
| Part 2: success rate (%) | 133 | 22% | 100% | 83% | 0.20 |
| Overall test success rate (%) | 133 | 30% | 100% | 78% | 0.19 |

The overall success rate was high in both parts of the test, which indicates a high level of knowledge results. Regarding the first part of the test, 53% of the respondents achieved an above average result (11 and more points out of 14), and 24% of the respondents achieved maximum number of points (14). In the second part of the test, 62% of the respondents achieved an above average result (8 or more points out of 9), 39% of them achieved maximum number of points (9).

The average success rate in the test was 78%, which is a very high value in terms of the general requirements for optimal test difficulty (around 50%), the success rate in the second part of the test (83%) even exceeds the maximum recommended value of 80% (Chráska, 2016). As early as this stage, we can state that the test is too easy for the given specific group of respondents, which will be apparent in the following parts of the analysis too.

### *Psychometric properties of individual items based on CTT*

Within the CTT we have focused on difficulty, discriminatory ability of items, and analysis of their distractors. We define difficulty as a percentage of respondents who have answered an item correctly (the higher the percentage, the easier the item is). This setting ensues from generally accepted parameterization of IRT models (Orlando and Thissen, 2020). Items answered correctly by more than 80% of the respondents are considered as very easy. Discriminatory ability is expressed in the form of so-called Upper-lower index (ULI), and discerns results between the worst and best thirds of the respondents. A quality item should, to a greater extent, favor students from the better group against those from the worse one. This index should reach at least the value of 0.2 (marked with a black line in the charts). However, this is not true for very difficult or very easy items. Those lack naturally good discernment. The results for both parts of the test are presented in Chart 1 and Chart 2.

Item 8 is the most difficult in Part 1 of the test (47% of respondents answered correctly). The easiest is Item 13 (97% correct answers). Almost all the items show low difficulty except Item 8. Items 6 and 13 have a low value of discrimination.

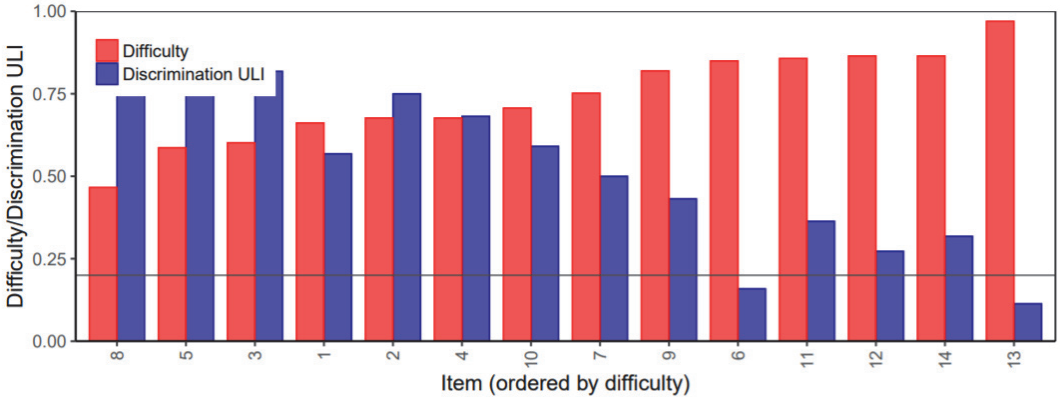


**Chart 1.** Difficulty and sensitivity of items in Part 1 of the test

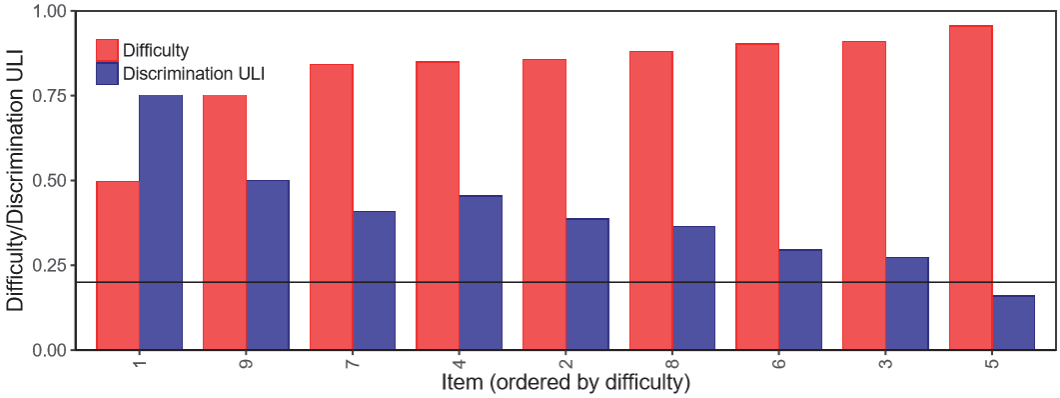


**Chart 1.** Difficulty and sensitivity of items in Part 2 of the test

Item 1 is the most difficult item in Part 2 (50% correct answers). Item 5 is the easiest (95% correct answers). All the items except for Item 1 can be considered as easy as more than 80% of the respondents answered them correctly. On top of that, Item 5 shows a low value of discrimination.

Within the analysis we reduced data to its binary form (true-false), thus not all the offered possibilities of distractors are reflected upon (2 to 3 in each item). There are distractors in the test that no one chose, namely in the following items in Part 1:

- 11. The best way to take care of your feet is to: b) massage them with alcohol each day.
- 12. Eating foods lower in fat decreases your risk of: d) eye disease.
- 13. Numbness and tingling may be symptoms of: c) eye disease.

Also, in the following items in Part 2:

- 3. If you have taken rapid-acting insulin, you are most likely to have a low blood glucose reaction in: c) 6–12 hours.
- 5. If you are beginning to have a low blood glucose reaction, you should: a) exercise.
- 6. A low blood glucose reaction may be caused by: d) too little exercise.
- 7. If you take your morning insulin but skip breakfast, your blood glucose level will usually: c) remain the same.

The above distractors appear to have been too easy in our study. In items 5 and 6, it is apparent that it is the same task only reversely formulated. To measure knowledge of nursing professionals, we recommend reviewing these items. However, it should be noted that these results are based on the application of the tool to nursing students prior to graduation, and that the tool was created in its original form for patients.

### *Psychometric properties of individual items based on the IRT*

While comparing models, we apply two indicators: AIC (Akaike information criteria) as well as BIC (Bayes information criteria), with the lowest value indicating the best model. Model comparison for both parts of the test is shown in Table 2.

In both cases, the lowest value of the BIC was achieved while applying a two-parameter model. This type of model reflects, besides the level of the respondents' knowledge, the difficulty and discriminatory ability of items (for a more detailed specification see Furr, 2018, pp. 462–463). The results for this model are presented below.

**Table 2. Model comparison for both parts of the test**

| AIC | BIC: Part 1 | BIC: Part 2 |
|---|---|---|
| 1 parametric model | 1728.029 | 859.605 |
| 2 parametric model | 1718.615 (lowest) | 851.661 (lowest) |
| 3 parametric model | 1722.49 | 869.457 |
| 4 parametric model | 1772.264 | 895.371 |

Chart 3 and Chart 4 (so-called characteristic item curves) express the development of probability of the correct answer (vertical axis) to individual items depending on the overall knowledge of the respondents (horizontal axis). Thus the *x* axis expresses the level of the respondents' knowledge that is standardized to a unified scale, from the weakest (–6), to average (0) to excellent (6). The knowledge is assessed based on the overall success rate of the respondents in the given part of the test. The *y* axis shows the probability of the correct answer to the given item from 0 to 1, where 1 expresses 100% probability of the correct answer. The placement of the items from left to right reflects their difficulty from the lowest to the highest. It is important that the curves are neither too flat nor too steep. It is also important that the curves are relatively evenly distributed from left to right around zero on the *x* axis.

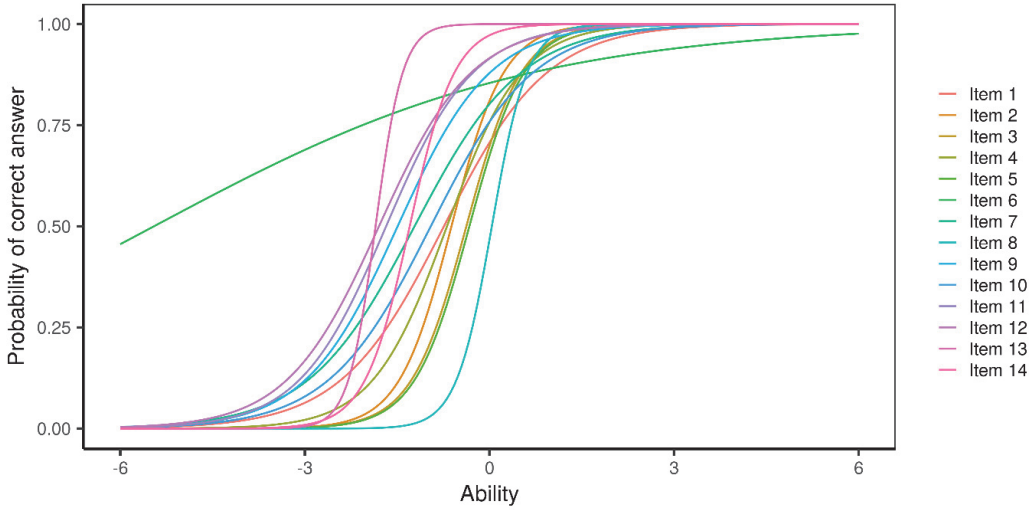


**Chart 3.** Probability of correct answer, Part 1

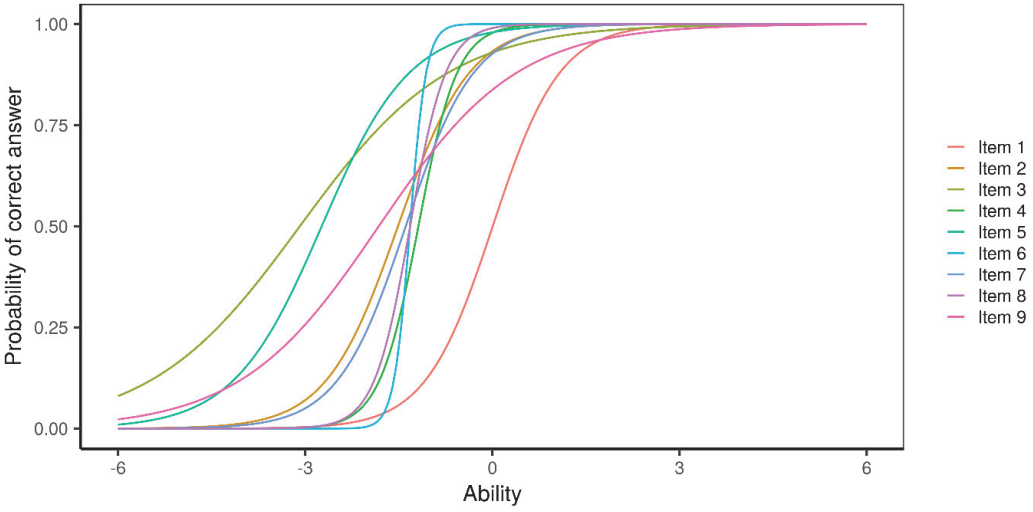


**Chart 4.** Probability of correct answer, Part 2

Chart 3, representing Part 1 of the test, shows the problem of Item 6. Its flattest green curve shows that the probability of answering this item correctly is only vaguely connected to the abilities of the respondents. This item has a weak discriminatory ability. However, for the test as a whole this is not a significant limitation. It is more serious when some items show very similar parameters (placement, gradient), which is better shown in Table 3. From left to right, the items are distributed relatively evenly. However, all of them are primarily placed in the left part of the chart, which shows a generally low difficulty of all the items in this test battery. Respondents with only a slightly above-average level of knowledge will very probably solve the whole test battery completely correctly.

Chart 4, representing Part 2 of the test, reflects the main problem again. The curves are mainly located in the left part of the chart, thus the items are generally easy. However, their stratification in the chart is relatively varied, which is desirable.

The next two charts express the overall characteristics of the two parts of the test, which is the relationship between the total information value of the test battery and knowledge, or the relationship between standard error and knowledge. In both cases, the test battery is beneficial for the detection of the level of knowledge in respondents with below-average knowledge (conclusions of Charts 3 and 4 are presented in Suppl. 1). With very low, and mainly above-average and increasing lev-

el of knowledge, the information curve drops significantly. Meanwhile, the curve expressing standard error is growing. Basically, the test cannot differentiate the level of knowledge when applied to respondents with very low, above-average, or excellent knowledge.

In the previous parts of this article, we criticized some items, and in others we found both negatives and positives. Another aid used to identify items that might be eliminated from the test are the values of Chi-squared test criteria (for more see Orlando and Thissen, 2000). *P*-values (see Table 3) lower than 0.05 indicate problematic items that are likely to be eliminated and replaced by others. These are the following items from Part 1 of the test:

- 4. Which of the following is a "free food"?
- 8. Which should not be used to treat low blood glucose?

**Table 3. Parameter estimates and item fit**

| | A | SE(a) | B | SE(b) | S-$\chi^2$ | DF | *P*–value |
|---|---|---|---|---|---|---|---|
| *Part 1* | | | | | | | |
| Item 1 | 1.19 | 0.32 | –0.74 | 0.23 | 2.34 | 7 | 0.939 |
| Item 2 | 2.35 | 0.58 | –0.62 | 0.14 | 5.93 | 5 | 0.313 |
| Item 3 | 2.23 | 0.52 | –0.38 | 0.14 | 1.39 | 5 | 0.925 |
| Item 4 | 1.65 | 0.41 | –0.69 | 0.18 | 17.31 | 6 | 0.008 |
| Item 5 | 2.23 | 0.52 | –0.33 | 0.14 | 1.29 | 5 | 0.936 |
| Item 6 | 0.32 | 0.29 | –5.45 | 4.68 | 3.98 | 7 | 0.782 |
| Item 7 | 1.15 | 0.34 | –1.22 | 0.31 | 8.73 | 6 | 0.189 |
| Item 8 | 3.46 | 0.90 | 0.04 | 0.12 | 11.88 | 3 | 0.008 |
| Item 9 | 1.33 | 0.40 | –1.49 | 0.35 | 7.83 | 6 | 0.251 |
| Item 10 | 1.20 | 0.33 | –0.95 | 0.26 | 5.86 | 7 | 0.556 |
| Item 11 | 1.41 | 0.46 | –1.68 | 0.39 | 0.68 | 5 | 0.984 |
| Item 12 | 1.33 | 0.47 | –1.80 | 0.46 | 5.06 | 5 | 0.409 |
| Item 13 | 4.51 | 3.31 | –1.85 | 0.36 | 0.00 | | |
| Item 14 | 2.74 | 0.94 | –1.30 | 0.21 | 7.12 | 3 | 0.068 |
| *Part 2* | | | | | | | |
| Item 1 | 1.85 | 0.55 | 0.01 | 0.15 | 1.51 | 1 | 0.219 |
| Item 2 | 1.74 | 0.52 | –1.51 | 0.30 | 3.46 | 3 | 0.326 |
| Item 3 | 0.84 | 0.40 | –3.08 | 1.26 | 3.19 | 3 | 0.363 |
| Item 4 | 3.20 | 1.10 | –1.18 | 0.18 | 0.27 | 2 | 0.874 |
| Item 5 | 1.41 | 0.62 | –2.73 | 0.83 | 2.10 | 1 | 0.147 |
| Item 6 | 7.71 | 6.41 | –1.32 | 0.17 | 0.00 | | |
| Item 7 | 1.82 | 0.53 | –1.39 | 0.27 | 1.60 | 3 | 0.659 |
| Item 8 | 3.48 | 1.22 | –1.32 | 0.19 | 2.16 | 2 | 0.340 |
| Item 9 | 0.90 | 0.32 | –1.82 | 0.57 | 5.94 | 3 | 0.115 |

*Note:* A = discrimination, SE(a) = standard error of discrimination, B = IRT difficulty (computed in 2-parameter IRT model), SE(b) = standard error of difficulty, S-$\chi^2$ = Chi-squared statistics (Orlando and Thissen, 2000), DF = degrees of freedom, *P*-value = statistical significance. The result of Chi-squared statistics is not available for items with the highest level of discrimination: Part A: Item 13, part B: Item 6 (Orlando and Thissen, 2000).

Item 4 seems to be problematic as it is of similar difficulty to Item 2 (Part 1) but has a significantly lower discrimination value. Item 8 is the most difficult of all the items, which makes it stand out among them. The results can also be compared to Chart 1. We consider this analysis as complementary, because the elimination of items should be conditioned mainly by expert assessment with regard to content validity of the test, and also with regard to the target group of respondents. Reducing a part of the test may appear to be a radical solution, and would require a repeated assessment of the psychometric parameters of the test as a whole. Of course, the elimination of some items is conditioned by giving some thought to the character of the test group, in our case students of the third year of nursing. For the general population, the behavior of the test would probably be different (more favorable).

### *Estimation of reliability*

The estimate of reliability is shown in Table 4. In the notion of the IRT (Furr, 2018, p. 477), test reliability varies depending on the given application (it will be higher in some respondents, and lower in others). The test shows higher reliability if there are respondents with various levels of knowledge represented in the set of respondents, because it is then easier to detect differences between the more and less knowledgeable ones. In relatively homogenous groups, a more sensitive test is needed to capture differences. Besides the traditionally used Cronbach's α coefficient, we also present the values of McDonald's ω and Guttman's λ6, thus trying to eliminate frequently criticized drawbacks of Cronbach's α. We have mentioned all the values including 95% of confidence intervals. In all cases, the value of more than 0.7 is reached. Both the test batteries

and the test as a whole can be considered as reliable within this application.

**Table 4. Estimate of test reliability including confidence intervals**

| Test part | Cronbach's α | McDonald's ω | Guttman's λ6 |
|---|---|---|---|
| Part 1 | 0.800 (0.746, 0.847) | 0.818 (0.773, 0.863) | 0.835 (0.815, 0.888) |
| Part 2 | 0.736 (0.664, 0.799) | 0,739 (0,673, 0.805) | 0.776 (0.708, 0.847) |
| Total | 0.858 (0.822, 0.888) | 0.868 (0.836, 0.901) | 0.906 (0.905, 0.943) |

## Discussion

Within our study, we applied the MDKT, namely its reviewed version (Fitzgerald et al., 2016). The test exists in several language versions. Most recent studies, for example, refer to the Arab version (Alhaiti et al., 2016; AlShayban et al., 2020; Madae'en et al., 2020), the Turkish version (İdiz et al., 2020), and the Norwegian version (Haugstvedt et al., 2016). The MDKT was also used to assess general knowledge of diabetes in a cross-sectional study by Ramjan et al. (2017). This study examined the connection between knowledge of and perception of diabetes mellitus (DM) among nursing students in Japan ($N$ = 78) and Australia ($N$ = 85), and at the same time compared differences in the respective curriculums.

Information of the validity of the test cannot be summarized using one coefficient as in the case of reliability. However, the test was created, repeatedly tested, and modified by experts in the field who assessed and discussed its content, which was relatively stable and well-attested while the test was being used.

Reliability of the various versions of the test (Alhaiti et al., 2016; AlShayban et al., 2020; Fitzgerald et al., 2016; Haugstvedt et al., 2016; İdiz et al., 2020; Madae'en et al., 2020) is in most cases acceptable, in some studies it is relatively weak (values below 0.7, as shown in Table 5). We found the lowest values in the case of a study performed on nursing staff (Haugstvedt et al., 2016) for whom the test was not originally intended.

**Table 5. Information of reliability of various tests**

| Study | Respondents | Reliability (Cronbach's alpha) | | | Language |
|---|---|---|---|---|---|
| | | Part A (1–14) | Part B (15–23) | Total | |
| İdiz et al. (2020) | 296 patients | 0.60 | 0.59 | 0.70 | Turkish |
| Kudlová and Kočvarová (2020) | 237 nurses (working in GPOs) | 0.68 | 0.52 | 0.74 | Czech |
| Beverly et al. (2019) | 230 medical students | 0.77 | 0.84 | – | English |
| Ramjan et al. (2017) | 78 students in Japan and 85 students in Australia | – | – | 0.80 | English |
| Alhaiti et al. (2016) | 139 patients | – | – | 0.75 | Arab |
| Haugstvedt et al. (2016) | 127 nursing staff | 0.57 | 0.42 | – | Norwegian |
| Fitzgerald et al. (2016) | 190 patients | 0.77 | 0.84 | – | English |
| Collins et al. (2011) | 99 patients | – | – | 0.71/0.61* | English |
| Fitzgerald et al. (1998) | 811 patients | 0.71 | 0.75 | – | English |

* Calculated for 2 versions of the test (1. simplified version, 2. revised version), GPOs – General Practitioners' Offices.

It is important to note that the results of reliability may be affected by the application of modified (not only in terms of language), versions of the test to different groups of respondents (nursing staff, patients). The problem that complicates comparison of study results is non-unified expression of reliability for individual parts of the test (which we consider as more accurate), or for the test as a whole (which we consider as disputable as it is well known that a higher number of items artificially increases overall reliability). Another problem is that authors of studies sometimes combine various research samples, e.g., Fitzgerald et al. (2016) presents reliability results for two interconnected sets of 101 and 89 respondents acquired in different ways, which may affect the results. The lowest reliability of individual parts of the test was seen in the study by Haugstvedt et al. (2016) conducted on nursing staff (Part A – 0.57, Part B – 0.42), and the highest reliability in a study by Fitzgerald et al. (2016) conducted on patients (Part A – 0.77, Part B – 0.84).

Our results are reliable, but in the context of this convenience sample, our results present the tool as less demanding for 3rd year students of nursing. Due to the sample characteristics and size, this outcome cannot be generalized. The authors therefore suggest extending the investigation to involve a larger population. The analysis points out the possibility of reducing the test, namely in the case of items 4 and 8 in Part 1. This reduction, however, is conditioned not only by statistical output, but primarily by expert judgment as to the content validity of the test. That is why we have left reduction as an open possibility that should also be considered with regard to target groups of respondents with further applications of the test.

It is necessary to point out that the acquired results are not influenced only by the quality of the test applied, but also by the respondents who filled it out. Using the test, we are able to measure their knowledge in the given area, however, not their courage to guess, effort paid to reading the task setting properly, etc.

### *Limitations of the study*

A limitation of the study is that the research set does not meet the requirements for representativeness because it was rea-

lized as a convenience sample. The research sample includes only students from one Czech university. That is why the results cannot be generalized for the whole population of Czech students of nursing. A further limitation is the relatively small sample size.

Another limitation concerns the application of the tool to nursing students. In our case, these were students at the end of their studies who were intensively preparing for their final state exams, thus their knowledge of the topic was fresh. Although the MDKT was originally created for patients, it is also commonly used around the world to test the knowledge of healthcare professionals and students. From our point of view, the tool originally intended for patients seems too simple for these students. Despite the above-mentioned limitations, the results can be considered unique in the Czech context.

Another step should be its application to nursing staff in clinical practice, and comparison to studies, in which this tool was applied to nursing staff. This is our next goal.

### *Implications for practice*

The diagnostic tool requires additional repeated measuring focused on its psychometric properties, which are not optimal. It is possible to consider shortening the tool while maintaining its psychometric qualities, or modifying problematic items. Before its application, it is always necessary to consider whether it is suitable for the given target group (especially to distinguish between patients and healthcare professionals, but also their previous experience in the given area).

## Conclusions

The test is reliable, but we do not consider it valid for our specific group of nursing students who are close to the end of their study (in the third year). The presented results primarily indicate that the test is too easy for our research sample. When the test is too easy, it cannot correctly differentiate the level of knowledge, which should be its main purpose. We suggest applying the MDKT at an earlier stage of the study, once the students have learned the basics of the issue and their knowledge is roughly at the level of that of educated patients. It is also possible to create and validate a more difficult version of the test, which could be used at the end of the study for students, but also for health care workers.

The test can also be reduced by the simplest or problematic items listed in the text. However, this could result in lowering its reliability and content validity (some areas would not be included in the test at all).

It is important to reiterate that our study is based on a convenient sample of third-year students of nursing from one Czech university. Therefore, the results cannot be generalized to a wider student population. It is possible that there is a specific approach to diabetes mellitus education at this university, which may have an impact on our research findings. Also, it is important to repeat that the MDKT was originally designed for patients. Other groups (students, medical staff) may need a (significantly) modified version, in terms of content and difficulty.

The question of validation and standardization of this test still seems open. Further studies on similar – as well as different samples – are needed to monitor the psychometric characteristics of the test under different circumstances.

### *Funding*

This paper is financed (translation, submission fees) by RVO excellent /MSM-RVO / FHS / 2021/005 as part of the project "Variable factors influencing health and disease". It is an internal grant within Tomas Bata University in Zlín designated for publishing research outcomes.

### *Ethical aspects and conflict of interests*

The authors have no conflict of interests to declare.

## Michiganský znalostní test o diabetu: ověření modifikované verze mezi českými studenty ošetřovatelství

**Souhrn**

Michigan Diabetes Knowledge Tool (MDKT) byl vyvinut s cílem otestovat úroveň znalostí v léčbě diabetu mellitu. Ve světě je široce používán jako originální (pro pacienty) i jako modifikovaný nástroj (pro sestry nebo studenty ošetřovatelství).

*Cíl:* Cílem je psychometrické testování modifikovaného MDKT.

*Metody:* Tato kvantitativní studie využívá deskriptivní statistiku, klasickou teorii testů (CTT), teorii odpovědi na položku (ITT) a odhad reliability v rámci dostupného výběru studentů ošetřovatelství ($N$ = 133). Sběr dat probíhal v letech 2018–2019 na vybrané české univerzitě.

*Výsledky:* Aplikace MDKT nám umožňuje rozlišovat pouze v rámci úzkého spektra studentů, jejichž znalosti jsou podprůměrné až průměrné. Analýza poukázala na možnost redukovat test při zachování jeho psychometrických parametrů.

*Závěr:* Test je reliabilní, nicméně příliš jednoduchý pro náš dostupný výběr studentů třetího ročníku ošetřovatelství, což však nelze zobecnit. Další aplikace testu mohou vést k jeho modifikacím, například vytvoření obtížnější verze testu a jeho opětovné ověření.

**Klíčová slova:** klasická teorie testu (CTT); Michiganský znalostní test o diabetu (MDKT); reliabilita; studenti ošetřovatelství; teorie odpovědi na položku (ITT); validita

## References


1. ADA – American Diabetes Association (2019). Introduction: Standards of medical care in diabetes. Diabetes Care 42(Suppl. 1): S1–S2. DOI: 10.2337/dc19-SINT01.
2. Alhaiti AH, Alotaibi AR, Jones LK, DaCosta C, Lenon GB (2016). Psychometric evaluation of the revised Michigan Diabetes Knowledge Test in Arabic: Translation and validation. J Diabetes Res 2016: 9643714. DOI: 10.1155/2016/9643714.
3. AlShayban DM, Naqvi AA, Alhumaid O, AlQahtani AS, Islam MA, Ghori SA, et al. (2020). Association of disease knowledge and medication adherence among out-patients with type 2 diabetes mellitus in Khobar, Saudi Arabia. Front Pharmacol 11: 60. DOI: 10.3389/fphar.2020.00060.
4. Bergenstal RM, Tamborlane WV, Ahmann A, Buse JB, Dailey G, Davis SN, et al. (2010). Effectiveness of sensor-augmented insulin-pump therapy in type 1 diabetes. N Engl J Med 363(4): 311–320. DOI: 10.1056/NEJMoa1002853.
5. Beverly EA, Ritholz MD, Rennie RG, Mort SC (2019). A brief interactive training with medical students improves their diabetes knowledge about hypoglycemia. BMC Med Educ 19(1): 171. DOI: 10.1186/s12909-019-1615-x.
6. CDS – Czech Diabetes Society (2020). Doporučený postup péče o diabetes mellitus 2. typu: Česká diabetologická společnost. Revize ze dne: 31. 5. 2020. Za ČDS: Škrha J, Pelikánová T, Prázný M, Kvapil M. [online] [cit. 2021-01-28]. Available at: https://www.diab.cz/dokumenty/Standardy_DM.pdf
7. Chan MF, Zang Y-L (2007). Nurses' perceived and actual level of diabetes mellitus knowledge: results of a cluster analysis. J Clin Nurs 16(7B): 234–242. DOI: 10.1111/j.1365-2702.2006.01761.x.
8. Chráska M (2016). Metody pedagogického výzkumu: základy kvantitativního výzkumu (2nd, updated ed). Praha: Grada, 256 p.
9. Christiansen MP, Klaff LJ, Brazg R, Chang AR, Levy CJ, Lam D, et al. (2018). A Prospective multicenter evaluation of the accuracy of a novel implanted continuous glucose sensor: PRECISE II. Diabetes Technol Ther 20(3): 197–206. DOI: 10.1089/dia.2017.0142.
10. Collins GS, Mughal S, Barnett AH, Fitzgerald JT, Lloyd CE (2011). Modification and validation of the revised diabetes knowledge scale. Diabet Med 28(3): 306–310. DOI: 10.1111/j.1464-5491.2010.03190.x.
11. Davies MJ, D'Alessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, et al. (2018). Management of hyperglycemia in type 2 diabetes. A Consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care 41(12): 2669–2701. DOI: 10.2337/dci18-0033.
12. Drass JA, Muir-Nash J, Boykin PC, Turek JM, Baker KL (1989). Perceived and actual level of knowledge of diabetes mellitus among nurses. Diabetes Care 12(5): 351–356. DOI: 10.2337/diacare.12.5.351.
13. Fitzgerald JT, Funnell MM, Anderson RM, Nwankwo R, Stantsfield RB, Piatt GA (2016). Validation of the revised brief diabetes knowledge test (DKT2). Diabetes Educ 42(2): 178–187. DOI: 10.1177/0145721715624968.
14. Fitzgerald JT, Funnell MM, Hess GE, Barr PA, Anderson RM, Hiss RG, Davis WK (1998). The reliability and validity of a brief diabetes knowledge test. Diabetes Care 21(5): 706–710. DOI: 10.2337/diacare.21.5.706.
15. Furr RM (2018). Psychometrics: an introduction. 3rd ed. Sage publications, Los Angeles, p. 543.
16. Haugstvedt A, Aarflot M, Igland J, Landbakk T, Graue M (2016). Diabetes knowledge in nursing homes and home-based care services: a validation study of the Michigan Diabetes Knowledge Test adapted for use among nursing personnel. BMC Nurs 15: 40. DOI: 10.1186/s12912-016-0159-1.
17. Hollis M, Glaister K, Lapsley JA (2014). Do practice nurses have the knowledge to provide diabetes self-management education? Contemp Nurse 46(2): 234–241. DOI: 10.5172/conu.2014.46.2.234.
18. İdiz C, Çelik S, Bağdemir E, Dişsiz M, Satman İ (2020). Turkish adaptation of michigan diabetes research and training center's revised diabetes knowledge test and determination of factors affecting the knowledge level of diabetic individuals [Diyabet araştırma ve eğitim merkezi revize diyabet bilgi testi'nin türkçeye adaptasyonu ve diyabetli bireylerin bilgi düzeyini etkileyen faktörlerin belirlenmesi]. Turk J Endocrinol Metab 24(1): 38–46. DOI: 10.25179/tjem.2019-71974.
19. Kudlová P, Kočvarová I (2020). Knowledge and self-assessment of general practitioner nurses in the area of diabetes mellitus. Kontakt 22(3): 165–171. DOI: 10.32725/kont.2020.031.
20. Kudlová P, Chlup R, Zapletalová J (2017). The benefits of postgraduate therapeutic education in diabetology in the Czech Republic. In 3rd ICHHPSY – International Conference on Health and Health Psychology, European Proceedings of Social & Behavioural Sciences 30: 217–227. DOI: 10.15405/epsbs.2017.09.20.
21. Madae'en S, Elayeh E, Akour A, Al Qhaiwi T, Shaggour B, Madain R (2020). Diabetes knowledge, medication adherence, and glycemic control among diabetic patients: A cross-sectional study in Jordan. J Appl Pharm Sci 10(4): 41–46. DOI: 10.7324/JAPS.2020.104007.
22. Mandysová P (2019). Metody překladu a transkulturní validace zahraničních nástrojů [Methods of translation and transcultural validation of foreign instruments]. 1st ed. Pardubice: Univerzita Pardubice, Fakulta zdravotnických studií, 129 p.
23. Martinková P, Drabinová A (2018). Shiny item analysis for teaching psychometrics and to enforce routine analysis of educational tests. R J 10(2): 503–515. DOI: 10.32614/RJ-2018-074.
24. Orlando M, Thissen D (2000). Likelihood-based item-fit indices for dichotomous item response theory models. Appl Psychol Meas 24(1): 50–64. DOI: 10.1177/01466216000241003.
25. Ramjan LM, Watanabe H, Salamonson Y (2017). Diabetes knowledge and perceptions among nursing students, and curriculum differences in Japan and Australia: A cross-sectional study. Nurse Educ Today 53: 7–12. DOI: 10.1016/j.nedt.2017.03.005.
26. Rhodes D, Visker J, Larson K, Cox C (2019). Rapid E-Learning for professional development in school-based diabetes management. Nurse Educ Pract 38: 84–88. DOI: 10.1016/j.nepr.2019.06.009.
27. Rubin DJ, Moshang J, Jabbour SA (2007). Diabetes knowledge: are resident physicians and nurses adequately prepared to manage diabetes? Endocr Pract 13(1): 17–21. DOI: 10.4158/EP.13.1.17.
28. Scheiderich SD, Freibaum CN, Peterson LM (1983). Registered nurses' knowledge about diabetes mellitus. Diabetes Care 6(1): 57–61. DOI: 10.2337/diacare.6.1.57.
29. World Health Organization (2020). Country and regional data on diabetes. [online] [cit. 2021-01-28]. Available from: https://www.who.int/news-room/fact-sheets/detail/diabetes
30. Yacoub MI, Demeh WM, Barr JL, Saleh AM, Saleh MY (2015). Outcomes of a diabetes education program for registered nurses caring for individuals with diabetes. J Contin Educ Nurs 46(3): 129–133. DOI: 10.3928/00220124-20150126-02.
31. Yacoub MI, Demeh WM, Darawad MW, Barr JL, Saleh AM, Saleh MY (2014). An assessment of diabetes-related knowledge among registered nurses working in hospitals in Jordan. Int Nurs Rev 61(2): 255–262. DOI: 10.1111/inr.12090.