

ANALIZA WYNIKÓW WSPOMAGANEGO KOMPUTEROWO ZALICZENIA TESTOWEGO Z PRZEDMIOTÓW „BADANIA NAUKOWE W PIELĘGNIARSTWIE” ORAZ „KSZTAŁCENIE W BADANIACH NAUKOWYCH W POŁOŻNICTWIE” DLA STUDENTÓW WYDZIAŁU NAUKI O ZDROWIU WARSZAWSKIEGO UNIwersYTETU MEDYCZNEGO

ANALYSIS OF RESULTS OF COMPUTER-AIDED FINAL TESTS FOR THE RESEARCH IN NURSING AND EDUCATION IN RESEARCH IN MIDWIFERY COURSES FOR STUDENTS OF FACULTY OF HEALTH SCIENCE, MEDICAL UNIVERSITY OF WARSAW

Joanna Gotlib¹, Mariusz Panczyk¹, Joanna Pawłowska², Marcin Malczyk²

¹ Zakład Dydaktyki i Efektów Kształcenia
Warszawski Uniwersytet Medyczny

² Biuro ds. Egzaminów Uczelnianych
Warszawski Uniwersytet Medyczny

DOI: <https://doi.org/10.20883/pielpol.2017.23>

STRESZCZENIE

Cel. Analiza porównawcza wyników zaliczenia testowego z przedmiotów „Badania naukowe w pielęgniarstwie” oraz „Kształcenie w badaniach naukowych w położnictwie” dla studentów WNoZ WUM.

Materiał i metody. Baza licząca 182 zadania testowe była przygotowana w formacie pytań wielokrotnego wyboru z jedną opcją prawidłową. Wszystkie pytania były indeksowane i poddawane parametryzacji za pomocą dwóch współczynników jakościowych – łatwość oraz moc różnicująca. W ocenie wyników uzyskanych z zaliczenia testowego przez studentów położnictwa i pielęgniarstwa posłużono się elementami statystyki opisowej oraz testami Levene’a, Grubbsa i Shapiro-Wilka. Zastosowano również analizy porównawcze: test t-Studenta dla powtórzonych pomiarów, test t-Studenta dla grup niezależnych oraz dwuczynnikową ANOVA. Wszystkie obliczenia wykonano z użyciem oprogramowania STATISTICA wersja 12.5. Dla wszystkich analiz jako domyślny (*a priori*) poziom istotności przyjęto $\alpha = 0,05$.

Wyniki. Stopień zróżnicowania punktacji zaliczeniowej dla obu przedmiotów był podobny (test Levena, $p=0,39$), nie zanotowano danych odstających (test Grubbsa, $p=1,00$), a częstość rozkładu w niewielkim stopniu była różna od rozkładu normalnego (test Shapiro-Wilka, $p=0,04$). Studenci położnictwa uzyskiwali wyższe wyniki niż pielęgniarstwa (test t-Studenta: $p=0,014$). Ponadto obserwowano spadkowy trend w zakresie uzyskiwanej przez studentów punktacji w kolejnych terminach zaliczenia (ANOVA: $p<0,001$). Te same zadania były trudniejsze dla studentów pielęgniarstwa (test t-Studenta: $p=0,035$), natomiast w zakresie mocy różnicującej zadania w grupie studentów pielęgniarstwa osiągały wyższe wartości. Liczba wyświetlań pytania nie wpływała istotnie na jego łatwość, niezależnie od analizowanego kierunku studiów.

ABSTRACT

Aim. Comparative analysis of the results of final tests for the *Research in Nursing and Education in Research in Midwifery* courses for students of the Faculty of Health Science, MUW.

Materials and methods. A total of 182 multiple-choice test questions with one correct option. Question parameters: easiness, differentiation power. Descriptive statistics as well as Levene's, Grubbs' and the Shapiro-Wilk tests. Comparative analysis: Student's t test for repeated measurements, Student's t test for independent groups, and two-factor ANOVA. STATISTICA version 12.5, $\alpha=0.05$.

Results. The degree of diversification of scores for both courses was similar (Levene's test, $p=0.39$), no outliers were found (Grubbs' test, $p=1.00$), and the frequency of distribution was slightly different from the normal distribution (the Shapiro-Wilk test, $p=0.04$). Midwifery students obtained better results compared to Nursing students (Student's t test: $p=0.014$). A declining trend in scores achieved by students at subsequent dates of the test was also observed (ANOVA: $p<0.001$). The same tasks were more difficult for Nursing students (Student's t test: $p=0.035$), whereas, as far as differentiation power is concerned, tasks reached higher values among Nursing students. The number of screen views of a question did not significantly influence its easiness, regardless of the major.

Wnioski

1. Analiza wyników zaliczenia z przedmiotów „Badania naukowe w pielęgniarstwie” oraz „Kształcenie w badaniach naukowych w położnictwie” wykazała, że wyniki uzyskane przez studentów położnictwa są istotnie wyższe niż pielęgniarstwa, co ze względu na fakt, że studenci obu kierunków studiów mieli taki sam program kształcenia, wymaga dalszych badań.
2. Możliwość przekazywania sobie przez studentów treści zadań testowych losowanych z bazy pytań w kolejnych turach oraz występowanie tych samych pytań w kolejnych turach zaliczenia nie wpływa istotnie na pogorszenie jakości oceniania.

SŁOWA KLUCZOWE: egzaminy wspomagane komputerowo, pomiar dydaktyczny, badania naukowe, studenci nauk o zdrowiu.

Conclusions

1. The analysis of the results of final tests for the *Research in Nursing and Education in Research in Midwifery* courses demonstrated that the scores achieved by Midwifery students were significantly better compared to those of Nursing students, which, due to the fact that both groups of students had the same curriculum, requires further study.
2. The possibility of exchanging information by students on test questions randomly selected from the database of questions in the subsequent rounds and the presence of the same questions in the subsequent rounds did not significantly influence the quality of assessment.

KEYWORDS: computer-aided exams, didactic measurement, research, Health Science students

Wstęp

Do końca roku akademickiego 2013/14 w Warszawskim Uniwersytecie Medycznym (WUM) egzaminy testowe prowadzone były wyłącznie w formie tradycyjnych egzaminów papierowych, z wykorzystaniem programu *Testy komputerowe* do analizy ich wyników [1, 2]. Od roku akademickiego 2014/15 istnieje w WUM możliwość przeprowadzania egzaminów testowych wspomaganych komputerowo, poprzez wykorzystanie platformy egzaminów elektronicznych eASK (ASK Systems) [3, 4].

W roku akademickim 2014/15 w WUM odbyło się 17 egzaminów wspomaganych komputerowo. Opinie studentów, którzy brali udział w tej formie egzaminowania są dobre [5], a studenci, którzy jeszcze nie brali udziału w egzaminach wspomaganych komputerowo również są pozytywnie nastawieni do tej nowatorskiej metody sprawdzania wiedzy [5]. Prowadzono również pilotażowe badania dotyczące opinii nauczycieli akademickich na temat tej formy egzaminowania, jednakże w tej grupie postawy wobec egzaminów wspomaganych komputerowo były bardzo zróżnicowane i skrajne. Konieczne jest jednak kontynuowanie badań w tym zakresie w liczniejszej grupie wykładowców [6].

Przedmioty „Badania naukowe w pielęgniarstwie” dla studentów III roku studiów pierwszego stopnia na kierunku pielęgniarstwo (w wymiarze 20 godzin wykładów i 10 godzin seminariów, semestr letni) oraz „Kształcenie w badaniach naukowych w położnictwie” dla studentów II roku studiów pierwszego stopnia na kierunku położnictwo (w wymiarze 30 godzin wykładów, semestr letni) prowadzone w Zakładzie Dydaktyki i Efektów Kształcenia Wydziału Nauki o Zdrowiu WUM odbywają się od wielu lat na platformie e-learningowej Moodle [7–12]. Kursy e-learningowe składają się z 6 modułów tematycznych: 1. Bazy piśmiennictwa naukowego, 2. Zasady cytowania i tworzenia bibliografii, 3. Formato-

wanie tekstu pracy dyplomowej do druku, 4. Rzetelność badań naukowych, 5. Bioetyczne aspekty prowadzenia badań naukowych w medycynie, 6. Polskie czasopisma naukowe. Oba przedmioty zakończone są zaliczeniem.

Zaliczenie przedmiotów w poprzednich latach odbywało się na platformie e-learningowej Moodle. Studenci zaliczali każdy moduł, który zakończony był testem lub zadaniem do wykonania. W roku akademickim 2014/15 zmodyfikowano formę zaliczenia i w czerwcu 2015 r. po raz pierwszy odbyło się ono w formie testu wspomagane komputerowo na platformie egzaminów eASK.

Cel pracy

Celem pracy była analiza wyników wspomagane komputerowo zaliczenia testowego z przedmiotów „Badania naukowe w pielęgniarstwie” oraz „Kształcenie w badaniach naukowych w położnictwie” dla studentów Wydziału Nauki o Zdrowiu WUM oraz porównanie uzyskanych przez studentów wyników zaliczenia w zależności od kierunku studiów

Materiał i metody

Organizacja zaliczenia

Biuro ds. Egzaminów Uczelnianych dysponuje salą wyposażoną w 30 stanowisk komputerowych przeznaczonych do zdawania przez studentów egzaminów i zaliczeń wspomaganych komputerowo, dlatego też odbywają się one zazwyczaj w kilku turach.

W zaliczeniu przedmiotów „Badania naukowe w pielęgniarstwie” oraz „Kształcenie w badaniach naukowych w położnictwie” wzięło udział łącznie 259 osób, a zaliczenie odbywało się w 12 turach. Szczegółową charakterystykę organizacji zaliczenia testowego z przedmiotów „Badania naukowe w pielęgniarstwie” oraz „Kształcenie w badaniach naukowych w położ-

nictwie” oraz studentów biorących udział w zaliczeniu testowym przedstawiono w **tabeli 1**.

Tabela 1. Organizacja zaliczenia testowego z przedmiotów „Badania naukowe w pielęgniarstwie” oraz „Kształcenie w badaniach naukowych w położnictwie” dla studentów pielęgniarstwa i położnictwa Wydziału Nauki o Zdrowiu WUM

Table 1. Arrangement of end-of-semester tests for the Research in Nursing and Training in Research in Obstetrics courses for Nursing and Midwifery students of the Faculty of Health Sciences, WUM

	Termin zaliczenia/ Dates of examinations	Pielęgniarstwo/ Nursing n=183	Położnictwo/ Obstetrics n=76	Łącznie/ Total n=259
1.	2015-06-01, godz. 8:30/ June 1st, 2015, 8:30 a.m.	15	7	22
2.	2015-06-01, godz. 12:30/ June 1st, 2015, 12:30 p.m.	10	17	27
3.	2015-06-02, godz. 8:30/ June 2nd, 2015, 8:30 a.m.	17	10	27
4.	2015-06-02, godz. 12:30/ June 2nd, 2015, 12:30 p.m.	15	13	28
5.	2015-06-10, godz. 8:30/ June 10th, 2015, 8:30 a.m.	13	15	28
6.	2015-06-10, godz. 12:30/ June 10th, 2015, 12:30 p.m.	24	4	28
7.	2015-06-10, godz. 15:15/ June 10th, 2015, 3:15 p.m.	25	0	25
8.	11/06/2015, godz. 15:00/ June 11th, 2015, 3:00 p.m.	0	3	3
9.	11/06/2015, godz. 18:00/ June 11th, 2015, 6:00 p.m.	9	6	15
10.	12/06/2015, godz. 08:30/ June 12nd, 2015, 8:30 a.m.	20	0	20
11.	12/06/2015, godz. 10:00/ June 12nd, 2015, 10:00 a.m.	6	0	6
12.	12/06/2015, godz. 13:30/ June 12nd, 2015, 1:30 p.m.	29	1	30

Źródło: opracowanie własne
Source: author's own analysis

W bazie pytań testowych eASK znalazły się 182 pytania. Pytania testowe losowane były przed każdą kolejną turą egzaminu. Każdy student przystępujący do zaliczenia testowego na platformie egzaminów elektronicznych eASK miał wylosowanych 30 pytań, każde z nich z 3 opcjami odpowiedzi i tylko jedną odpowiedzią prawidłową. Próg zaliczenia przedmiotu został ustalony na 17 punktów (56%).

Analiza jakości zadań testowych

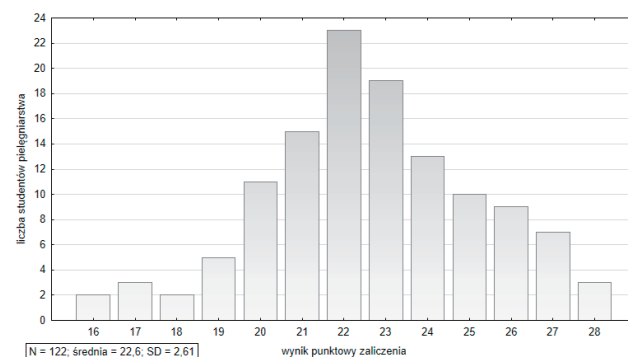
Baza licząca 182 zadania testowe była przygotowana w formie pytań wielokrotnego wyboru z jedną opcją prawidłową. W toku zaliczenia na obu kierunkach studiów powtórzyły się łącznie 154 zadania (84,6%). Wszystkie pytania były indeksowane i poddane parametryzacji za pomocą dwóch współczynników jakościowych – łatwość oraz moc różnicująca.

W ocenie wyników uzyskanych z zaliczenia testowego przez studentów położnictwa i pielęgniarstwa posłużono się elementami statystyki opisowej oraz testami Levene’a (jednorodność wariancji), Grubbsa (obecność danych odstających) i Shapiro-Wilka (zgodność z rozkładem normalnym). Natomiast do analizy porównawczej przebiegu zaliczenia testowego dla obu kierunków studiów wykorzystano parametryczny test t-Studenta dla powtórzonych pomiarów, test t-Studenta dla grup niezależnych oraz dwuczynnikową analizę wariancji (two-way ANOVA). Dodatkowo, dla każdego porównania wyznaczono także miarę efektu za pomocą współczynnika *d*-Cohena oraz cząstkowe eta-kwadrat (η^2). Do oceny potencjalnej zależności między łatwością pytania testowego a liczbą jego wyświetleń podczas zaliczenia wyznaczono nieparametryczny współczynnik korelacji *r*-Spearmana.

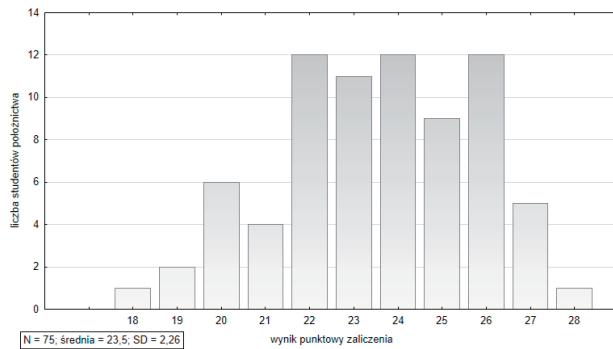
Wszystkie obliczenia wykonano z użyciem oprogramowania STATISTICA wersja 12.5 (StatSoft, Inc.), zgodnie z licencją WUM. Dla wszystkich analiz jako domyślny (*a priori*) poziom istotności przyjęto $\alpha=0,05$.

Wyniki

Analizując wyniki uzyskane przez studentów pod kątem rozkładu tej zmiennej można zaobserwować, że stopień zróżnicowania poszczególnych wyników był podobny (test Levene, $p=0,39$), nie zanotowano obecności danych odstających (test Grubbsa, $p=1,00$), a częstość rozkładu była tylko w niewielkim stopniu różna od rozkładu normalnego (test Shapiro-Wilka, $p=0,04$). Rozkład wyników był nieco ujemnie skośny, co jest typowe dla tego typu danych. Szczegółowe dane dotyczące struktury punktacji uzyskanej na zaliczeniu testowym przez studentów obu kierunków przedstawiono na **rycynie 1**.



Liczba studentów pielęgniarstwa/Number of Nursing students; Wynik punktowy zaliczenia/Credit score



Liczba studentów położnictwa/Number of Obstetrics students; Wynik punktowy zaliczenia/Credit score

Rycina 1. Wyniki statystyki opisowej oraz częstości rozkładu punktacji uzyskanej na zaliczeniu testowym przez studentów pielęgniarstwa i położnictwa

Figure 1. Results of descriptive statistics and distribution of frequency of scores obtained for the test by Nursing and Obstetrics students

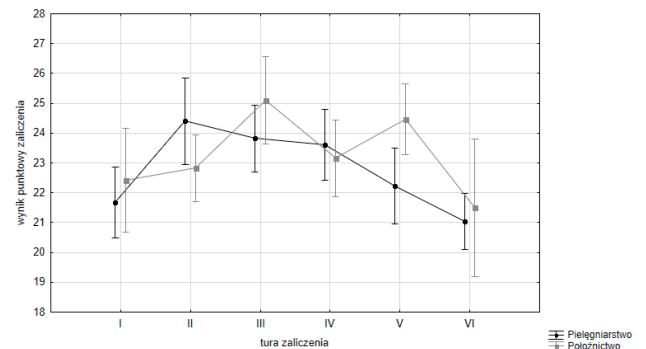
Średnia/mean; SD – odchylenie standardowe/standard deviation

Źródło: opracowanie własne
Source: author's own analysis

Porównując punktację otrzymaną przez studentów obu kierunków studiów zaobserwowano, że studenci położnictwa uzyskiwali istotnie statystycznie wyższe wyniki niż studenci pielęgniarstwa (test t-Studenta: $t=-2,477$; $p=0,014$). Wartość współczynnika d -Cohena dla tego porównania wyniosła 0,37.

Oszacowując wpływ kolejnych tur zaliczenia oraz kierunku studiów na uzyskane przez studentów wyniki testu, posłużono się dwuczynnikową analizą wariancji. Niezależnie od kierunku studiów wyniki zaliczenia w kolejnych terminach różniły się między sobą istotnie (ANOVA: $F=4,51$; $p<0,001$; cząstkowe $\eta^2=0,13$). Ponadto zaobserwowano, że istnieje spadkowy trend w zakresie punktacji w kolejnych terminach zaliczenia, jednakże jego intensywność była silniej zaznaczona w przypadku studentów pielęgniarstwa (interakcja kierunek studiów*tura zaliczenia; ANOVA: $F=2,15$, $p=0,062$; cząstkowe $\eta^2=0,07$). Szczegółowe wyniki analizy wariancji przedstawiono na **rycynie 2**.

Łatwość poszczególnych zadań testowych była zróżnicowana, a te same zadania były istotnie statystycznie trudniejsze dla grupy studentów pielęgniarstwa w porównaniu ze studentami kierunku położnictwa (test t-Studenta: $t=-2,135$; $p=0,035$). Wartość współczynnika d -Cohena dla tego porównania wyniosła 0,14. Uzyskane wyniki porównania łatwości poszczególnych pytań testowych dla dwóch kierunków studiów przedstawiono na **rycynie 3**.

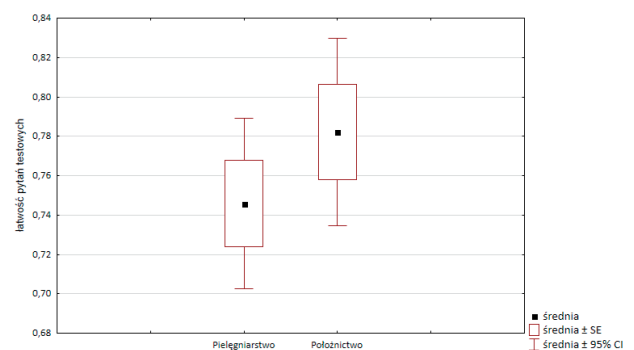


Pielęgniarstwo/Nursing, Położnictwo/ Obstetrics; Tura zaliczenia/round of end-of-semester test

Rycina 2. Wpływ kolejnych tur zaliczenia oraz kierunku studiów na wyniki uzyskane w toku zaliczenia (opis wykresu: centralny punkt – średnia, wąsy – 95% przedział ufności)

Figure 2. Influence of subsequent rounds of the end-of-semester test and the study major on scores obtained for the test (figure description: central points – mean, whiskers – 95% confidence interval)

Źródło: opracowanie własne
Source: author's own analysis



Łatwość pytań testowych/Easiness of questions; Pielęgniarstwo/Nursing; Położnictwo/Obstetrics

Średnia/mean; SE – błąd standardowy/standard deviation; 95% CI – 95% przedział ufności/confidence interval

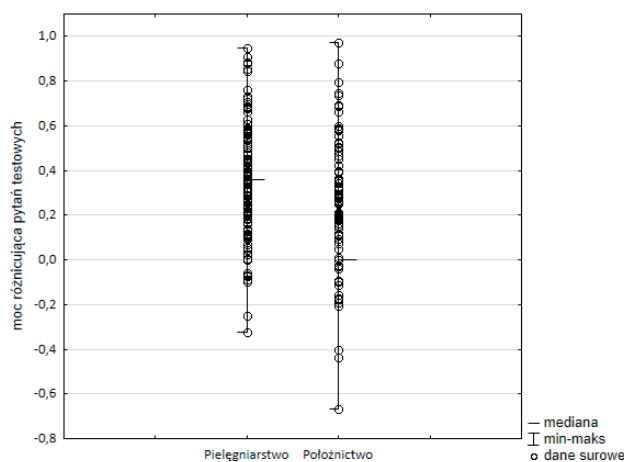
Rycina 3. Porównanie wartości parametru łatwości dla zadań testowych użytych podczas zaliczenia dla dwóch kierunków studiów (schemat porównania dla grup powiązanych, czyli powtórzonych pomiarów)

Figure 3. Comparison of easiness of test questions used in the end-of-semester test for both majors (comparison scheme for related groups, i.e. repeated measures)

Źródło: opracowanie własne
Source: author's own analysis

Również pod względem parametru mocy różnicującej zaznaczała się różnica między dwoma kierunkami studiów. W przypadku zaliczenia dla studentów pielęgniarstwa te same pytania miały wyższe wartości mocy różnicującej niż dla położnictwa (**Rycina 4**). Jednakże różnica ta ma związek z różną liczebnością obu porównywanych grup studentów, co ma istotny wpływ na

obserwowaną odmienną zdolność różnicującą tych samych pytań w dwóch różnych grupach studenckich.



Moc różnicująca pytań testowych/*Discriminating power of test questions*; Pielęgniarstwo/*Nursing*; Położnictwo/*Obstetrics*; Mediana/min-maks/dane surowe/*Median/min.-max./raw data*

Rycina 4. Porównanie parametru mocy różnicującej pytań testowych użytych podczas zaliczenia w dwóch grupach studentów (schemat porównania dla grup powiązanych, czyli powtórzonych pomiarów)

Figure 4. Comparison of discriminating power of test questions used in the end-of-semester test among two groups of students (comparison scheme for related groups, i.e. repeated measures)

Źródło: opracowanie własne
Source: author's own analysis

Analizując wpływ liczby wyświetlań danego pytania na jego łatwość, nie zaobserwowano istotnych statystycznie korelacji, niezależnie od ocenianego kierunku studiów. Zestawienie wyników analizy korelacji przedstawiono w **tabeli 2**.

Tabela 2. Wyniki analizy korelacji między łatwością danego pytania a liczbą jego wyświetleń podczas zaliczenia dla dwóch kierunków studiów
Table 2. Results of analysis of correlation between easiness of a particular question and a number of its views during the test in relation to both majors

Kierunek studiów/ <i>Field of study</i>	Liczba pytań/ <i>number of questions</i>	Korelacja r-Spearmana/ <i>r-Spearman's correlation</i>	Wartość statystyki t/ <i>a value of the t-statistic</i>	P
Pielęgniarstwo/ <i>Nursing</i>	154	-0,082	-1,013	0,313
Położnictwo/ <i>Obstetrics</i>	154	-0,065	-0,796	0,427

Źródło: opracowanie własne
Source: author's own analysis

Dyskusja

W dostępnym, polskim piśmiennictwie naukowym (Polska Bibliografia Lekarska, słowa kluczowe: pielęgniar-

ki – nauczanie, położne – nauczanie, 2000–2015) nie odnaleziono dotychczas publikacji dotyczących szczegółowych analiz jakości narzędzi egzaminacyjnych, w szczególności egzaminów i zaliczeń w formie testów wykorzystywanych do sprawdzania wiedzy studentów, oraz pośrednio do analizy skuteczności i efektywności kształcenia. Analizy jakości narzędzi egzaminacyjnych – egzaminów i zaliczeń w formie testów, zarówno w wersji tradycyjnej, jak i wspomaganą komputerowo – opisywane są w literaturze dotychczas tylko przez zespół autorów niniejszej publikacji [13–20].

Opisywane zagadnienie jest w dostępnym, krajowym piśmiennictwie tematem nowatorskim. Tematyka jakości narzędzi egzaminacyjnych nie była dotychczas podejmowana. Dotychczas publikowane prace z obszaru oceny jakości kształcenia na kierunkach medycznych i związanych z naukami o zdrowiu dotyczą częściej kształcenia praktycznego lub satysfakcji studentów ze studiowania. Doświadczenia uczelni związane z egzaminami wspomaganymi komputerowo również nie były dotychczas opisywane w polskim piśmiennictwie naukowym, być może ze względu na fakt, że obecnie w Polsce tylko trzy uniwersytety medyczne wykorzystują platformy egzaminów elektronicznych do sprawdzania wiedzy studentów: Uniwersytet Medyczny w Poznaniu, Uniwersytet Medyczny w Łodzi oraz Warszawski Uniwersytet Medyczny.

Prezentowana w niniejszej pracy analiza wyników wspomaganego komputerowo zaliczenia testowego z przedmiotów „Badania naukowe w pielęgniarstwie” oraz „Kształcenie w badaniach naukowych w położnictwie” dla studentów Wydziału Nauki o Zdrowiu WUM jest pierwszą w polskim piśmiennictwie z obszaru nauk o zdrowiu publikacją dotyczącą analizy jakości pytań zawartych w teście końcowym z przedmiotu dotyczącego badań naukowych.

Analizując wyniki zaliczenia z przedmiotów „Badania naukowe w pielęgniarstwie” oraz „Kształcenie w badaniach naukowych w położnictwie” uzyskane przez studentów pielęgniarstwa i położnictwa Wydziału Nauki o Zdrowiu WUM można zaobserwować, że stopień zróżnicowania poszczególnych wyników był podobny, nie zanotowano obecności danych odstających, a częstość rozkładu była w niewielkim stopniu różna od rozkładu normalnego. Uzyskane wyniki zaliczenia przeprowadzonego na platformie egzaminów elektronicznych eASK nie odbiegają od wyników uzyskiwanych przez studentów podczas tradycyjnych zaliczeń testowych, odbywających się w tradycyjnej formie papierowej. Zdaniem autorów, przeprowadzenie zaliczeń w formie testów wspomaganą komputerowo jest zatem równie skuteczne i efektywne, co zaliczeń odbywających się w tradycyjnej formie, ponadto mają one

wiele innych zalet, których nie mają egzaminy papierowe, takich jak możliwość zastosowania pytań multimedialnych, uprzedniających zaliczenia testowe, czy szybkość otrzymywanych wyników [13–20].

Celem prezentowanej pracy było porównanie wyników uzyskanych podczas zaliczenia wspomaganego komputerowo przez studentów pielęgniarstwa i położnictwa.

Porównując punktację otrzymaną przez studentów obu kierunków studiów zaobserwowano, że studenci położnictwa uzyskiwali wyższe wyniki niż studenci pielęgniarstwa. Łatwość poszczególnych zadań testowych była zróżnicowana, a te same zadania były trudniejsze dla grupy studentów pielęgniarstwa w porównaniu ze studentami położnictwa. Również pod względem parametru mocy różnicującej zaznaczała się różnica między dwoma kierunkami studiów. W przypadku zaliczenia dla studentów pielęgniarstwa te same pytania miały wyższe wartości mocy różnicującej niż dla położnictwa. Jednakże różnica ta ma związek z różną liczebnością obu porównywanych grup studentów, co ma istotny wpływ na obserwowaną odmienną zdolność różnicującą tych samych pytań w dwóch różnych grupach studenckich. Wyniki zaliczenia w kolejnych terminach różniły się między sobą istotnie, ponadto zaobserwowano, że istnieje spadkowy trend w zakresie punktacji w kolejnych terminach zaliczenia, jednakże jego intensywność była silniej zaznaczona w przypadku studentów pielęgniarstwa.

Uzyskane wyniki wymagają, zdaniem autorów, dalszych badań. Studenci obu kierunków studiów brali udział w tych samych kursach e-learningowych na platformie egzaminów elektronicznych eASK, dlatego też zdaniem autorów zaskakujące są różnice w wynikach uzyskanych przez studentów pielęgniarstwa i położnictwa. Jednym z powodów tego zjawiska może być różnica w liczebności studentów poszczególnych kierunków studiów. W związku z tym, postuluje się, aby uzyskane wyniki potraktować jako wyniki badania pilotażowego i kontynuować analizy w kolejnych latach.

Jednym z najważniejszych zagadnień branych pod uwagę podczas planowania egzaminów wspomaganego komputerowo jest budowa bazy pytań testowych, a w szczególności liczba pytań, które powinny znaleźć się w bazie, aby można było przeprowadzić tego typu egzamin. W światowych wytycznych dotyczących egzaminów wspomaganego komputerowo nie odnaleziono jednoznacznych zaleceń dotyczących konkretnej liczby pytań, które powinny znaleźć się w bazie pytań testowych [21]. Liczba pytań powinna zależeć zarówno od treści kształcenia z wybranego przedmiotu, jak również od planu testu: łącznej liczby pytań oraz planowanego procentu wymiany pytań w poszczególnych turach egzaminu, jak również od liczby studentów biorących

udział w egzaminie czy też liczby stanowisk komputerowych, którymi dysponuje uczelnia. Zdaniem autorów niniejszej pracy, bardzo istotnym zagadnieniem jest zbudowanie odpowiedniej bazy pytań równoważnych, posiadających odpowiednie parametry łatwości (pytania łatwe, średnie i trudne) oraz mocy różnicującej. Zbudowanie takiej bazy pytań jest warunkiem koniecznym do przeprowadzenia egzaminu testowego o pożądanej jakości [13–20].

Podczas organizacji egzaminu wspomaganego komputerowo istnieje możliwość losowania poszczególnych zadań z bazy pytań testowych. Jest to również jedna z najczęściej zgłaszanych przez nauczycieli akademickich obaw dotyczących przeprowadzania zaliczeń wspomaganego komputerowo, szczególnie w przypadku, kiedy zaliczenie organizowane jest w sesji egzaminacyjnej ciągłej i trwa nawet kilka dni lub tygodni. Nauczyciele obawiają się wzajemnego przekazywania sobie pytań testowych przez studentów po kolejnych turach egzaminów i powstawania tzw. „giełdy pytań”. Zaprezentowane w niniejszej pracy wyniki analizy pokazały, że punktacja uzyskana przez studentów w kolejnych terminach nie tylko nie była wyższa, lecz istnieje w tym zakresie trend spadkowy. Ponadto analizując wpływ liczby wyświetlań danego pytania na jego łatwość nie zaobserwowano istotnych statystycznie korelacji, niezależnie od ocenianego kierunku studiów. Nieuzasadnione zatem wydają się prezentowane często obawy nauczycieli akademickich związane z przekazywaniem sobie przez studentów treści pytań testowych losowanych z bazy w kolejnych turach egzaminu oraz występowaniem tych samych pytań w kolejnych turach zaliczenia.

Prezentowane w niniejszej pracy wyniki analiz wspomaganego komputerowo zaliczenia testowego wskazują zatem, że jest to dobra, akceptowana przez studentów forma zaliczenia, której wyniki są zbliżone do tradycyjnych zaliczeń testowych, jednakże prawidłowy przebieg zaliczeń wspomaganego komputerowo wymaga przygotowania bazy równoważnych pytań testowych odpowiedniej jakości, czyli o zweryfikowanych parametrach łatwości i mocy różnicującej.

Wnioski

1. Analiza wyników wspomaganego komputerowo zaliczenia testowego z przedmiotów „Badania naukowe w pielęgniarstwie” oraz „Kształcenie w badaniach naukowych w położnictwie” wykazała, że wyniki uzyskane przez studentów położnictwa są istotnie wyższe niż wyniki studentów pielęgniarstwa, co ze względu na fakt, że

studenci obu kierunków studiów mieli taki sam program kształcenia, wymaga dalszych badań.

2. Niezależnie od kierunku studiów wyniki zaliczenia w kolejnych terminach różniły się między sobą istotnie – zaobserwowano, że istnieje spadkowy trend w zakresie punktacji w kolejnych terminach zaliczenia. Nieuzasadnione zatem wydają się obawy związane z przekazywaniem przez studentów treści pytań testowych losowanych z bazy pytań w kolejnych turach egzaminu oraz występowaniem tych samych pytań w kolejnych turach zaliczenia.

Piśmiennictwo

1. System opracowania egzaminów testowych „Testy komputerowe” i „Ankiety” – <http://www.testy.com.pl/index.html> (data dostępu 10.05.2016).
2. Regulamin studiów na Warszawskim Uniwersytecie Medycznym wprowadzony na mocy uchwały nr 26/2013 Senatu WUM z dnia 22 kwietnia 2013 r.
3. Zarządzenie nr 33/2014 Rektora Warszawskiego Uniwersytetu Medycznego w sprawie powołania Biura ds. Egzaminów Elektronicznych z dnia 22 maja 2014 r.
4. Schemat organizacyjny Warszawskiego Uniwersytetu Medycznego – Załącznik nr 2 do Zarządzenia nr 33/2014 Rektora Warszawskiego Uniwersytetu Medycznego w sprawie powołania Biura ds. Egzaminów Elektronicznych z dnia 22 maja 2014 r.
5. Gotlib J, Panczyk M, Gębski P, Zarzeka A, Iwanow L, Dąbrowski F, Dykowska G, Malczyk M. Comparative analysis of opinions of Warsaw Medical University students on electronic examinations and final tests by their participation in this form of knowledge assessment – preliminary report. *Pol J Public Health* 2015; 3: 153–158.
6. Gotlib J, Panczyk M, Gębski P, Zarzeka A, Iwanow L, Dąbrowski F, Dykowska G, Malczyk M. Porównanie opinii studentów i nauczycieli Warszawskiego Uniwersytetu Medycznego na temat egzaminów i zaliczeń wspomaganých komputerowo – doniesienie wstępne. *Zdr Publ* 2015; 4: 89–94.
7. Platforma e-learningowa Moodle Zakładu Dydaktyki i Efektów Kształcenia Wydziału Nauki o Zdrowiu Warszawskiego Uniwersytetu Medycznego <http://www.nzd.moodle.wum.edu.pl/> (data dostępu 7.12.2015 r.).
8. Gotlib J, Białoszewski D, Mosiołek A, Dąbrowski F. Opinie studentów i nauczycieli akademickich kierunków medycznych na temat kształcenia przez Internet na przykładzie Warszawskiego Uniwersytetu Medycznego. Badanie pilotażowe. *Zdr Publ* 2012;122(1):13–19.
9. Gotlib J. Opinie studentów pielęgniarstwa na temat realizacji zajęć dydaktycznych z przedmiotu „Prawo w ochronie zdrowia” w formie kursu e-learningowego. *Med Dydak Wychow* 2012; XLIV (1): 19–23.
10. Gotlib J, Białoszewski D, Mosiołek A, Dąbrowski F. Analiza opinii studentów i nauczycieli akademickich Warszawskiego Uniwersytetu Medycznego na temat kształcenia przez Internet. *Med Dydak Wychow* 2012; XLIV (1): 25–32.
11. Panczyk M, Woynarowska-Sołdan M, Belowska J, Zarzeka A, Gotlib J. Analiza porównawcza przebiegu i efektywności kształcenia e-learningowego studentów pielęgniarstwa w zakresie „Rzetelności w badaniach naukowych”. Międzynarodowa Konferencja Naukowo-Szkoleniowa „Innowacyjność w kształceniu, praktyce pielęgniarstwa i badaniach naukowych. 90-lecie powstania Uniwersyteckiej Szkoły Pielęgniarek i Opiekunek Zdrowia w Krakowie”. 11–12 września 2015 r. Kraków, Polska. Streszczenia prac, 20–21.
12. Panczyk M, Belowska J, Gotlib J. Computer-aided testing: assessment of automatic item generation to create multiple choice test items. ICERI 2014, The 7th International Conference of Education, Research and Innovation – 17th-19th November 2014, Seville, Spain. ICERI2014 Proceedings 7th International Conference of Education, Research and Innovation November 17th-19th, 2014-Seville, Spain. Edited by L. Gómez Chova, A. López Martínez, I. Candel Torres. IATED Academy ISBN: 978-84-617-2484-0. ISSN: 2340-1095. Depósito Legal: V-2632-2014, 4139–4147.
13. Gotlib J, Grochans E, Panczyk M, Jurczak A, Stanisławska M, Malczyk M. Analiza wyników egzaminów testowych z przedmiotu „Podstawy pielęgniarstwa” jako wstęp do standaryzacji egzaminu testowego – badanie wielośrodkowe. *Doniesienie wstępne. Piel Pol* 2016; 3: 287–291.
14. Gotlib J, Zarzeka A, Panczyk M, Malczyk M. Zaliczenie testowe z przedmiotu „Prawo w medycynie” dla studentów Wydziału Nauki o Zdrowiu na platformie egzaminów elektronicznych ASK Systems – doświadczenia własne. *Medycyna Dydaktyka Wychowanie* 2015; 1: 28–31.
15. Gotlib J, Grochans E, Jurczak A, Stanisławska M, Panczyk M. Analiza wyników egzaminów z „Podstawy pielęgniarstwa” jako wstęp do standaryzacji egzaminu testowego – badanie wielośrodkowe. *Doniesienie wstępne. Międzynarodowa Konferencja Naukowo-Szkoleniowa „Innowacyjność w kształceniu, praktyce pielęgniarstwa i badaniach naukowych. 90-lecie powstania Uniwersyteckiej Szkoły Pielęgniarek i Opiekunek Zdrowia w Krakowie”. 11–12 września 2015 r. Kraków, Polska. Streszczenia prac, 24.*
16. Gotlib J, Zarzeka A, Panczyk M, Gębski P, Iwanow L, Malczyk M, Belowska J. Analiza jakości pytań z końcowego testu zaliczeniowego z przedmiotu „Prawo w medycynie” dla studentów pielęgniarstwa na platformie egzaminów elektronicznych ASK Systems. Międzynarodowa Konferencja Naukowo-Szkoleniowa „Innowacyjność w kształceniu, praktyce pielęgniarstwa i badaniach naukowych. 90-lecie powstania Uniwersyteckiej Szkoły Pielęgniarek i Opiekunek Zdrowia w Krakowie”. 11-12 września 2015 r. Kraków, Polska. Streszczenia prac, 74–75.
17. Gotlib J, Panczyk M, Gębski P, Olędzka G, Belowska J, Zarzeka A, Malczyk M. Analiza jakości egzaminacyjnych pytań testowych z przedmiotu „Mikrobiologia” dla studentów pielęgniarstwa jako metoda przygotowania bazy pytań do egzaminu wspomaganego komputerowo. Międzynarodowa Konferencja Naukowo-Szkoleniowa „Innowacyjność w kształceniu, praktyce pielęgniarstwa i badaniach naukowych. 90-lecie powstania Uniwersyteckiej Szkoły Pielęgniarek i Opiekunek Zdrowia w Krakowie”. 11–12 września 2015 r. Kraków, Polska. Streszczenia prac, 147–148.
18. Gotlib J, Panczyk M, Olędzka G, Belowska J, Zarzeka A, Malczyk M. Doskonalenie bazy testowych pytań egzaminacyjnych z przedmiotu „Mikrobiologia” dla studentów Pielęgniarstwa jako metoda podnoszenia jakości egzaminu wspomaganego komputerowo. Międzynarodowa Konferencja Naukowo-Szkoleniowa „Jubileusz 40-lecia Wydziału Nauk o Zdrowiu Uniwersytetu Medycznego im. Karola Marcinkowskiego w Poznaniu”. 26–27 czerwca 2015 r. Poznań, Polska. Zeszyt streszczeń, 99.
19. Gotlib J, Panczyk M, Gębski P, Dąbrowski F, Gałązkowski R, Malczyk M, Jabiry-Zieniewicz Z, Belowska J, Zarzeka A. Modification of the base of test exam questions from the „Gynaecology and Obstetrics” course for emergency medi-

- cine students as a form of improvement of an e-exam. Proceedings of EDULEARN15 Conference 6th-8th July 2015, Barcelona, Spain. ISBN: 978-84-606-8243-1. 0188–0197.
20. Gotlib J, Panczyk M, Zarzeka A, Dąbrowski F, Gębski P, Dykowska G, Malczyk M. Analysis of history of granting credits and handling test exams on the ask systems e-exam platform at Warsaw Medical University – a preliminary report. Proceedings of EDULEARN15 Conference 6th-8th July 2015, Barcelona, Spain. ISBN: 978-84-606-8243-1. 0198–0205.
21. Dennick R, Wilkinson S, Purcell N. Online e-Assessment: AMEE Guide No. 39 Medical Teacher 2009; 31: 192-206.
- Artykuł przyjęty do redakcji: 13.06.2016
Artykuł przyjęty do publikacji: 17.08.2016

Źródło finansowania: Praca nie jest finansowana z żadnego źródła.
Konflikt interesów: Autorzy deklarują brak konfliktu interesów.

Adres do korespondencji:

Joanna Gotlib
ul. Żwirki i Wigury 61
02-091 Warszawa
tel.: 22 5720 490, fax: 22 5720 491
e-mail: joanna.gotlib@wum.edu.pl
Zakład Dydaktyki i Efektów Kształcenia
Warszawski Uniwersytet Medyczny