

ANNA PAWŁOWSKA¹, TOMASZ ZDROJEWSKI², JACEK WIERZBOWSKI¹,
HANNA JASIEL-WOJCULEWICZ², PIOTR BANDOSZ², KRYSZTOF ADRYCH¹,
KAMIL CHWOJNICKI², MARIAN SMO CZYŃSKI¹, ZBIGNIEW ZDROJEWSKI³,
BOGDAN WYRZYKOWSKI²

**ZASTOSOWANIE TESTÓW DLA OCENY WIEDZY Z PROGRAMU
NAUCZANIA CHORÓB WEWNĘTRZNYCH NA 3 I 4 ROKU WYDZIAŁU
LEKARSKIEGO AKADEMII MEDYCZNEJ W GDAŃSKU**

APPLYING TESTS FOR THE VALUATION OF THE STUDENTS'
KNOWLEDGE OF INTERNAL MEDICINE ACQUIRED
IN 3RD AND 4TH YEAR OF STUDYING AT MEDICAL DIVISION
OF MEDICAL UNIVERSITY IN GDAŃSK

¹Klinika Gastroenterologii i Hepatologii AM w Gdańsku
kierownik: dr hab. Marian Smoczyński, prof. nzw.

²Katedra Nadciśnienia Tętniczego i Diabetologii
kierownik: prof. dr Bogdan Wyrzykowski

³Klinika Nefrologii, Transplantologii i Chorób Wewnętrznych
kierownik: prof. dr Bolesław Rutkowski

Testy stają się obecnie coraz popularniejszą formą sprawdzania wiedzy na uczelniach medycznych. Wyniki testów zależą nie tylko od przygotowania studentów, ale także od jakości pytań, z których składa się test. Artykuł przedstawia wyniki, strukturę i cechy testów z interny, użytych do kontroli wiedzy nabytej na 3 i 4 roku studiów medycznych. Zawiera także przykładowe pytania o różnej łatwości i mocy różnicującej. Monitorowanie parametrów testu, takich jak łatwość, moc różnicująca i rzetelność, jest istotne w kontroli jakości testu i ułatwia dostosowanie pytań do potrzeb dydaktycznych.

Na studiach medycznych w ostatnich latach obserwuje się coraz częstsze wykorzystywanie testów do sprawdzania stopnia opanowania materiału przez studentów [1]. Studenci sami często postulują wprowadzenie formy testowej egzaminów ze względu na jednakowy dla wszystkich stopień trudności i zakres sprawdzanej wiedzy oraz mniejszy udział przypadku przy uzyskiwaniu oceny pozytywnej lub negatywnej, jak to może mieć miejsce przy egzaminach ustnych.

Formę testu ma również wprowadzony niedawno Lekarski Egzamin Państwowy, zdawany przez absolwentów akademii medycznych po ukończeniu stażu podyplomowego.

Taka forma egzaminów umożliwia sprawdzenie nie tylko znajomości faktów, ale również rozumienia związków przyczynowych i analizowania danych [2]. Umożliwiają to różne rodzaje pytań testowych: pytania z wyborem prostym, wyborem wielokrotnym, z zaprzeczeniem, pytania typu: przesłanka–teza. Test składa się z wielu pytań, co umożliwia dość szczegółową ocenę znacznego zakresu wiadomości, a niezajomość niewielkiej części materiału nie powinna znacząco wpłynąć na wynik testu.

Forma testowa sprawdzianów jest szczególnie przydatna przy egzaminowaniu dużej ilości osób, gdyż oszczędza czas potrzebny na sprawdzanie odpowiedzi i umożliwia szybkie podanie wyników.

Testy nie są z pewnością idealną formą kontroli wiedzy i nie sprawdzają się w każdych warunkach. Są pozbawione możliwości indywidualnej oceny toku rozumowania studenta, radzenia sobie z problemami klinicznymi, rozważania różnych scenariuszy sytuacji klinicznych [3, 4]. Sprawdzają wiedzę bierną (wybór odpowiedzi) i nie wymagają czynnego przedstawienia informacji. Należy podkreślić, że rezygnacja z formy ustnej i praktycznej egzaminu byłaby błędem. Opieranie się jedynie na wynikach egzaminu testowego nie pozwala w sposób rzetelny ocenić wiadomości studentów. Rozumienie pytania i przedstawienie toku myślenia umożliwia egzaminatorowi wyciągnięcie wniosków dotyczących rzeczywistej wiedzy pytanego.

Istotną wadą testów jest duży nakład pracy konieczny do przygotowania puli dobrej jakości pytań testowych.

Ważne jest monitorowanie jakości testów. Parametry, które mówią o jakości pytań testowych to: współczynnik łatwości, moc różnicująca i rzetelność testu [5]. Łatwość podawana jest jako odsetek osób, które prawidłowo odpowiedziały na dane pytanie. Duża ilość pytań bardzo łatwych lub bardzo trudnych znacznie zmniejsza wartość testu. Moc różnicująca to parametr określający, jak dane pytanie odróżnia osoby, które generalnie dobrze rozwiązały cały test, od tych, które rozwiązały go słabo. Wartość zbliżona do jedności świadczy o wysokiej mocy różnicującej, zbliżona do zera lub ujemna – o niskiej. Rzetelność to parametr informujący o teście jako całości. Przedstawia stosunek różnic systematycznych i losowych w wynikach osób rozwiązujących test. Wartość współczynnika rzetelności powyżej 0,7 oznacza, że różnice uzyskane w punktacji przez poszczególnych studentów nie są przypadkowe, lecz obrazują realną wiedzę.

Celem artykułu jest przedstawienie wyników testów przeprowadzanych wśród studentów 5 roku Wydziału Lekarskiego, sprawdzających wiedzę nabytą na 3 i 4 roku z zakresu chorób wewnętrznych.

MATERIAŁ I METODY

Na początku bloku ćwiczeń z chorób wewnętrznych studenci rozwiązywali test, obejmujący elementy interny, nauczane na 3 i 4 roku. Składał się on z dwóch części: pierwsza dotyczyła propedeutyki chorób wewnętrznych, druga – wybranej dziedziny interny (nefrologia, kardiologia, hematologia, gastroenterologia). Zaliczenie testu z propedeutyki interny było warunkiem dopuszczenia do ćwiczeń z chorób wewnętrznych.

Badaniami objęto 391 studentów 5 roku WL. W ciągu 2 lat sprawdzian testowy przeprowadzono 10 razy. Test z propedeutyki interny składał się ze 100 pytań. Każde pytanie było zbudowane z 5 dystraktorów. Test należało rozwiązać w ciągu 100 minut.

Strukturę i cechy przykładowego testu opisuje tabela I.

Tab. I

Struktura i cechy testów

The structure and the features of the tests

Test / Test	Propedeutyka interny Propaedeutics of Internal Medicine	Test z przedmiotu Department test
Ilość pytań / Number of questions	100	50
Pytania proste / Simple questions	71	41
Pytania złożone / Composed questions	26	9
Pytania z ilustracją graficzną Graphic questions	3	0
Pytania zbyt łatwe Too easy questions	29	16
Pytania o średnim stopniu trudności Questions of intermediate difficulty	67	29
Pytania zbyt trudne Too difficult questions	4	5

Tab. II

Parametry testów z propedeutyki interny uzyskane w kolejnych testowanych grupach

The parameters of tests in the Propaedeutics of Internal Medicine in the consecutive groups

Numer grupy Group's number	Ilość zdających Number of students	Ilość ocen negatywnych Number of negative marks	Średnia punktów (%) Average score	Rzetelność testu Reliability of the test
1	40	7	79	0,82
2	40	3	80	0,75
3	40	7	79	0,69
4	37	5	79	0,83
5	37	6	75	0,73
6	46	5	80	0,85
7	45	7	79	0,83
8	34	5	74	0,73
9	35	10	74	0,75
10	37	11	74	0,77

WYNIKI

Poszczególne grupy uzyskały średnią punktów od 74% do 80% (do zaliczenia wymagano udzielenia 70% prawidłowych odpowiedzi). W sumie 66 studentów (17%) nie zaliczyło testu, w każdej grupie od 7,5 do 29,7% zdających. Współczynnik rzetelności testu (oceniany w każdej grupie zdających) był wysoki i wynosił od 0,69 do 0,85 (średnio 0,78). Parametry uzyskane we wszystkich grupach przedstawia tabela II.

Test z poszczególnych dziedzin interny składał się z 50 pytań, jego zaliczenie nie było warunkiem dopuszczenia do ćwiczeń z chorób wewnętrznych. Studenci nie byli wcześniej informowani, jakiej dziedziny będzie dotyczył test. Przeprowadzono testy: trzy razy z hematologii, dwa razy z nefrologii, dwa razy z gastroenterologii i trzy razy z kardiologii. W poszczególnych testowanych grupach studenci uzyskali punktację od 56% do 70%, średnio 63,3%. Rzetelność testu była dość zróżnicowana: od 0,48 do 0,93 (średnio 0,68). Parametry poszczególnych testów przedstawione są w tabeli III.

Tab. III

Parametry testów z poszczególnych dziedzin interny w kolejnych testowanych grupach

The parameters of tests in the particular departments of internal medicine in the consecutive groups

Numer grupy Group number	Przedmiot / Subject	Średnia punktów (%) Average score (%)	Rzetelność testu Reliability of the test
1	hematologia / hematology	64	0,87
2	nefrologia / nephrology	70	0,66
3	gastroenterologia / gastroenterology	61	0,70
4	kardiologia / cardiology	56	0,71
5	kardiologia / cardiology	60	0,52
6	kardiologia / cardiology	56	0,48
7	gastroenterologia / gastroenterology	62	0,52
8	nefrologia / nephrology	70	0,61
9	hematologia / hematology	68	0,78
10	hematologia / hematology	66	0,93

DYSKUSJA

Przedstawiony materiał zawiera wyniki testów z chorób wewnętrznych przeprowadzonych w ciągu dwóch lat na 5 roku studiów Wydziału Lekarskiego, przed rozpoczęciem ćwiczeń z chorób wewnętrznych (sprawdzian dopuszczający).

Analiza parametrów łatwości i mocy różnicującej poszczególnych pytań umożliwia usuwanie z bazy pytań o niskiej wartości (zbyt łatwych, zbyt trudnych i o niskiej mocy różnicującej)

oraz jest pomocą przy układaniu nowych testów. Omawiane testy cechowały się dość dużą rzetelnością.

Wynik testu zależy od wiedzy studenta, ale istotne znaczenie ma również charakterystyka pytań testowych: ich łatwość i moc różnicująca. Przykładowe pytania testowe z poszczególnych dziedzin interny podajemy w celu zapoznania czytelników z ich parametrami (aneks).

Czytelników zainteresowanych przykładowymi pytaniami z diagnostyki ogólnej o różnym stopniu trudności i różnej mocy różnicującej odsyłamy do artykułów zamieszczonych w „Polskim Archiwum Medycyny Wewnętrznej” [6, 7].

Dobrze ułożony test powinien zawierać dużo pytań o dobrej mocy różnicującej i jak najmniej pytań zbyt łatwych i zbyt trudnych, o niskiej (lub zerowej) mocy różnicującej.

Sprawdzian testowy służył dwóm celom: oceniał stopień przygotowania studentów do ćwiczeń w klinikach chorób wewnętrznych oraz umożliwił nauczycielom akademickim zapoznanie się z efektywnością nauczania interny na 3 roku (propedeutyka) i na 4 roku (nauczanie głównych dziedzin chorób wewnętrznych). Zaliczenie sprawdzianu z propedeutyki interny było obowiązkowe przed rozpoczęciem ćwiczeń i studenci mogli się do tego egzaminu przygotować. Zaliczenie drugiego testu nie było konieczne, jednak w badanych grupach wszyscy studenci pisali również ten test. Brak negatywnych konsekwencji słabego wyniku testu oraz niemożność przygotowania się do niego były prawdopodobnie przyczynami gorszych wyników tego testu niż testu z propedeutyki interny. Przeprowadzane testy miały również na celu systematyczne przygotowanie studentów do egzaminu testowego z chorób wewnętrznych (po 5 roku) i do Państwowego Egzaminu Lekarskiego.

WNIOSKI

1. Test ma swoje ograniczenia i nie zastępuje egzaminu ustnego.
2. Test jest cennym narzędziem jako egzamin przesiewowy, ponieważ umożliwia w krótkim czasie sprawdzenie znajomości obszernych zakresów materiału u dużej liczby osób.
3. Monitorowanie parametrów testu, takich jak łatwość, moc różnicująca i rzetelność jest istotne w kontroli jakości testu i znacznie ułatwia korektę bazy pytań, tak, by optymalnie odpowiadała ona potrzebom dydaktycznym.

PIŚMIENNICTWO

1. Mavis B. E., Cole B. L., Hoppe R. B.: A survey of information sources used for progress decisions about medical students. *Med. Educ. Online* 2000, 5, 9. – 2. Case S. M., Swanson D. B.: Constructing written test questions for the basic and clinical sciences. 3 ed. National Board of Medical Examiners, Philadelphia, PA 19104 (online) http://www.nbme.org/PDF/ItemWriting_2003/2003IWGwhole.pdf. – 3. Słomkiewicz S.: Egzamin w wyższej uczelni. *Życ. Szk. Wyższ.* 1983, 5, 27. – 4. Ballantyne C.: The Teaching and Learning Centre of Murdoch University (online): Multiple Choice Tests (2002). – 5. Niemierko B.: Testy osiągnięć szkolnych: podstawowe pojęcia i techniki obliczeniowe. Warszawa: WSiP, 1975. – 6. Zdrojewski T., Jasiel-Wojculewicz H.: Doświadczenia z wprowadzaniem testów do oceny nauczania propedeutyki interny na III roku Wydziału Lekarskiego Akademii Medycznej w Gdańsku. Cz. 1. *Pol. Arch. Med. Wewn.* (w druku). – 7. Jasiel-Wojculewicz H., Zdrojewski T.: Zastosowanie testów w ocenie nauczania propedeutyki interny: 10 lat doświadczeń własnych w Akademii Medycznej w Gdańsku. Cz. 2. *Pol. Arch. Med. Wewn.* (w druku).

A. Pawłowska, T. Zdrojewski, J. Wierzbowski, H. Jasiel-Wojculewicz,
P. Bandosz, K. Adrych, K. Chwojncki, M. Smoczyński, Z. Zdrojewski, B. Wyrzykowski

APPLYING TESTS FOR THE VALUATION OF THE STUDENTS' KNOWLEDGE
OF INTERNAL MEDICINE ACQUIRED IN 3RD AND 4TH YEAR OF STUDYING
AT MEDICAL DIVISION OF MEDICAL UNIVERSITY IN GDAŃSK

Summary

Test becomes presently more and more popular for controlling students' knowledge in medical universities. The results of the tests depend not only on students' adequate preparation, but also on the quality of the test questions.

The article contains the description of results, structure and features of the tests in internal medicine, used to control the knowledge of elements of internal medicine taught in 3rd and 4th year of study. The article contains also examples of questions of various difficulty and discrimination index. Monitoring of the parameters of the test: difficulty, discrimination index and reliability is very important to control the quality of the test and enables adjustment of the tests to the teaching needs.

Adres: Anna Pawłowska
Klinika Gastroenterologii i Hepatologii AMG
ul. Dębinki 7, Gdańsk
e-mail: apawłowska@amg.gda.pl

ANEKS

Przykłady pytań

Pytanie o dużym współczynniku łatwości (> 95%) i niskiej mocy różnicującej**Nefrologia:**

Najbardziej przydatnym i najłatwiej dostępnym badaniem nieinwazyjnym u chorego z podejrzeniem kamicy układu moczowego jest:

- A. tomografia komputerowa
- B. urografia
- C. ultrasonografia
- D. scyntygrafia nerek
- E. rezonans magnetyczny

Hematologia:

Powiększenie węzłów chłonnych śródpiersia spotyka się często w:

- A. przewlekłej białaczce szpikowej
- B. chorobie Hodgkina
- C. szpiczaku mnogim
- D. czerwienicy prawdziwej
- E. ostrej białaczce szpikowej

Kardiologia:

Za niestabilność blaszki miażdżycowej, predysponującej do jej pęknięcia odpowiadają:

- A. nadciśnienie tętnicze
- B. cienka torebka włóknista
- C. nacieki komórkowe w obrębie zmienionej miażdżycowo ściany tętniczej
- D. duże stężenie cholesterolu w obrębie struktury miażdżycowej
- E. wszystkie wymienione

Gastroenterologia:

Przyczyną cholestazy zewnątrzwątrobowej nie jest:

- A. kamica żółciowa przewodowa
- B. guzy głowy trzustki
- C. guz brodawki Vatera
- D. pierwotna marskość żółciowa
- E. jatrogenne (pooperacyjne) uszkodzenie dróg żółciowych

Pytanie bardzo trudne (współczynnik łatwości <15%), o niskiej mocy różnicującej**Nefrologia:**

W przebiegu którego zespołu/ choroby nie występuje zespół nerczycowy:

- A. toczeń rumieniowaty układowy
- B. amyloidoza
- C. cukrzyca
- D. zwyrodnienie wielotorbielowate nerek
- E. rak oskrzelowopochodny

Hematologia:

Objawy ogólne, towarzyszące zespołom limfoproliferacyjnym to:

1. poty
2. stany podgorączkowe i gorączki
3. osłabienie
4. chudnięcie
5. brak apetytu

Prawdziwe:

- A. wszystkie prawdziwe
- B. 1, 2, 5
- C. 1, 2, 4
- D. wszystkie nieprawdziwe
- E. 1, 2, 3

Kardiologia:

Wskaż prawidłową sekwencję zjawisk podczas niedokrwienia:

- A. ból – upośledzenie relaksacji – upośledzenie kurczliwości – zmiany w ekg
- B. ból – upośledzenie kurczliwości – upośledzenie relaksacji – zmiany w ekg
- C. zmiany w ekg – ból – upośledzenie relaksacji – upośledzenie kurczliwości
- D. upośledzenie relaksacji – upośledzenie kurczliwości – zmiany w ekg – ból
- E. upośledzenie relaksacji – upośledzenie kurczliwości – ból – zmiany w ekg

Gastroenterologia:

Zapalenie przełyku jest:

- A. jedną z istotnych przyczyn zaburzeń wchłaniania
- B. stwierdzane u prawie 100% chorych z refluksem żołądkowo-przełykowym
- C. wywoływane zakażeniem bakterią *Helicobacter pylori*
- D. leczone głównie lekami z grupy blokerów pompy protonowej
- E. występuje zawsze w skojarzeniu z przepukliną rozworu przełykowego

Pytanie o średnim stopniu trudności i dobrej mocy różnicującej**Nefrologia:**

Mikroskopowy krwimocz nie jest charakterystyczny dla:

- A. nefropatii IgA
- B. kamicy układu moczowego
- C. gruźlicy nerek
- D. nefropatii cukrzycowej
- E. urazu nerki

Hematologia:

Spośród wymienionych:

1. w hemofilii A typowe są wylewy dostawowe i domięśniowe
 2. w leczeniu hemofilii A należy podać czynnik VIII domięśniowo
 3. hemofilia A jest najczęstszą wrodzoną skazą osoczną
 4. w diagnostyce hemofilii A należy oznaczyć poziom czynnika VIII
- A. tylko 1, 2 są prawdziwe
 - B. tylko 2, 3 są prawdziwe
 - C. tylko 3, 4 są prawdziwe
 - D. tylko 1, 4 są prawdziwe
 - E. wszystkie są prawdziwe

Kardiologia:

U chorego z utratami przytomności w przebiegu całkowitego bloku przedsionkowego nie należy wszczepiać stymulatora:

- A. VVI
- B. VVIR
- C. AAI
- D. DDD
- E. VDD

Gastroenterologia:

Wybierz fałszywe stwierdzenie dotyczące wrzodziejącego zapalenia jelita grubego:

- A. najczęstsza lokalizacja to końcowy odcinek jelita krętego
- B. zmiany zapalne ograniczają się do błony śluzowej, rzadziej do podśluzowej
- C. przedmiotowo rzadko stwierdza się występowanie oporu w jamie brzusznej
- D. około połowa chorych zgłasza w przebiegu choroby dolegliwości stawowe
- E. leczenie chirurgiczne polega na resekcji jelita grubego.