

# Dobra praktyka w zakresie zapewniania jakości walidacji

## Wielka Brytania

*Kwalifikacja: Lekarz medycyny*

Warszawa 2018

Dodana: 2018-01-11



**Fundusze Europejskie**  
Wiedza Edukacja Rozwój

IBE



*kwalifikacje dla każdego*

**Unia Europejska**  
Europejski Fundusz Społeczny



*Dobra praktyka w zakresie zapewniania jakości walidacji. Wielka Brytania. Kwalifikacja: Lekarz medycyny wewnętrznej*

Na podstawie ekspertyzy autorstwa Katarzyny Ludki-Stempień

Redakcja merytoryczna:  
Sylwia Walicka, Iwona Gmaj

Redakcja językowa:  
Karolina Kwiatosz

Konsultacja merytoryczna:  
dr Agnieszka Chłoń-Domińczak

Skład i łamanie:  
dr Marcin Trepczyński

Instytut Badań Edukacyjnych  
ul. Górczewska 8  
01-180 Warszawa  
tel. (22) 241 71 00; [www.ibe.edu.pl](http://www.ibe.edu.pl)

Wzór cytowania:  
Instytut Badań Edukacyjnych (2017). Dobra praktyka w zakresie walidacji efektów uczenia się.  
Wielka Brytania. Kwalifikacja: Lekarz medycyny wewnętrznej. Warszawa: Instytut Badań Edukacyjnych.

© Copyright by: Instytut Badań Edukacyjnych, Warszawa 2017  
Publikacja współfinansowana przez Unię Europejską ze środków Europejskiego Funduszu Społecznego  
w ramach projektu Wspieranie realizacji I etapu wdrażania Zintegrowanego Systemu  
Kwalifikacji na poziomie administracji centralnej oraz instytucji nadających kwalifikacje i zapewniających  
jakość nadawania kwalifikacji.

Egzemplarz bezpłatny

## Spis treści

1. Kraj, nazwa instytucji . . . . .	4
2. Strona internetowa instytucji . . . . .	4
3. Nazwa kwalifikacji, której dotyczy dobra praktyka . . . . .	4
4. Dobra praktyka - Zewnętrzne zapewnianie jakości. . . . .	4
4.1. Informacje wstępne . . . . .	4
4.2. Obecny kształt egzaminu MRCP(UK). . . . .	4
4.2.1. Część I egzaminu . . . . .	5
4.2.2. Część II egzaminu . . . . .	8
4.2.3. Praktyczna ocena umiejętności klinicznych (PACES) . . . . .	10
4.2.4. Rzetelność i trafność wyników egzaminu . . . . .	14
4.3. Wieki doświadczenia prowadzące do MRCP(UK). . . . .	16
4.3.1. The Royal College of Physicians of London . . . . .	16
4.3.2. Ustawa z 1858 roku i jej konsekwencje. . . . .	17
4.3.3. Początki XX wieku. . . . .	18
4.3.4. Lata sześćdziesiąte XX wieku i wzrost naboru na studia wyższe. . . . .	18
4.3.5. Od lat siedemdziesiątych XX wieku do dnia dzisiejszego. . . . .	19
4.3.5.1. Format pytań i punktacja . . . . .	19
4.3.5.2. Ocena psychometryczna. . . . .	20
4.3.5.3. Wyodrębnienie PACES oraz badanie jego trafności i rzetelności . . . . .	21
5. Ograniczenia/Uwagi . . . . .	22

## 1. Kraj, nazwa instytucji

- Wielka Brytania
- The Royal College of Physicians of London (RCP)

## 2. Strona internetowa instytucji

<https://www.rcplondon.ac.uk/>

## 3. Nazwa kwalifikacji, której dotyczy dobra praktyka

The Examination for the Membership of the Royal Colleges of Physicians of the United Kingdom – MRCP(UK)

## 4. Dobra praktyka - Zewnętrzne zapewnianie jakości

### 4.1. Informacje wstępne

Dobra praktyka dotyczy zapewniania jakości egzaminu MRCP(UK), czyli obowiązkowego egzaminu wymaganego na ścieżce kariery lekarskiej w Wielkiej Brytanii, organizowanego przez trzy college'e lekarskie w Londynie, Edynburgu i Glasgow (Royal Colleges of Physicians, RCP). College'e są instytucjami, które na mocy edyktu królewskiego lub ustanowionego przez parlament aktu prawnego uzyskały prawo do sprawowania kontroli nad programem nauczania w określonej dziedzinie wiedzy. Pełnią zatem funkcję korporacji zawodowych.

Pełna nazwa egzaminu MRCP(UK) to The Examination for the Membership of the Royal Colleges of Physicians of the United Kingdom. Angielskie słowo *physician*, które w wolnym tłumaczeniu określa po prostu lekarza, w Wielkiej Brytanii oznacza specjalistę medycyny wewnętrznej.

MRCP(UK) jest jednym z egzaminów lekarskich o charakterze międzynarodowym, organizowanym na szeroką skalę. Co roku przystępuje do niego około 6000 osób z całego świata, z czego zaledwie 30% to absolwenci uczelni brytyjskich. Egzamin odbywa się, poza Wielką Brytanią, w 20 innych krajach, między innymi w: Indiach, Zjednoczonych Emiratach Arabskich, Egipcie, na Sri Lance. MRCP(UK) jest nie tylko egzaminem o światowej renomie, lecz także dobrym przykładem rzetelnego i trafnego narzędzia pomiarowego, którego historia pokazuje, jak zmiany społeczne i rozwój wiedzy wpływają na proces walidacji i nadawania kwalifikacji.

Dobra praktyka dotyczy dwóch aspektów:

- funkcjonujących obecnie rozwiązań w zakresie zapewniania jakości walidacji dla kwalifikacji specjalisty medycyny wewnętrznej;
- utrwalonego przez kilka wieków doskonalenia opisu tej kwalifikacji oraz procesu walidacji prowadzącego do formalnego potwierdzenia jej uzyskania.

Należy zaznaczyć, że rozwiązania dotyczące opisu kwalifikacji MRCP(UK) oraz egzaminu prowadzącego do potwierdzenia jej uzyskania są w sposób ciągły doskonalone, a prezentowana dobra praktyka zawiera opis rozwiązań według stanu na koniec 2016 r.

### 4.2. Obecny kształt egzaminu MRCP(UK)

Egzamin MRCP(UK) został opracowany w sposób umożliwiający wytypowanie najlepszych kandydatów na lekarzy prezentujących odpowiednią wiedzę, umiejętności i zachowania, które w ocenie college'u umożliwią im ukończenie specjalizacji. Zdanie MRCP(UK) gwarantuje dostęp do kontynuacji nauki na następujących specjalizacjach: dermatologia, endokrynologia, gastroenterologia, geriatria, choroby zakaźne, onkologia, nefrologia, neurologia, medycyna paliatywna, pulmonologia oraz reumatologia. Royal Colleges of Physicians (RCP) organizuje także, we współpracy z Brytyjskim Towarzystwem Kardiologicznym, specjalizację z kardiologii. Egzamin MRCP(UK) jest również wymagany w niektórych college'ach specjalizacyjnych np. w The Royal College of Radiologists (radiologia).

Egzamin składa się z trzech części:

- część I: test pisemny,
- część II: test pisemny,
- część III: egzamin kliniczny PACES (*Practical Assessment of Clinical Examination Skills*)

Egzamin odbywa się w trzech sesjach (odpowiadających poszczególnym częściom) organizowanych **w różnych miesiącach roku**. Przed rokiem 2009 należało przystępować do poszczególnych części egzaminu zgodnie z podaną wyżej kolejnością, natomiast obecnie można najpierw zdać część III: PACES, a dopiero później przystąpić do części II. Część I nadal zdaje się na samym początku. Bieżące informacje dotyczące egzaminu, przykładowe pytania, zasady przystępowania, porady oraz wszystkie dodatkowe informacje dla kandydatów dostępne są na stronie internetowej :[www.mrcpuk.org](http://www.mrcpuk.org). Wynik egzaminu kandydaci otrzymują w formie elektronicznej, po zalogowaniu się do systemu.

Kandydaci na lekarzy zazwyczaj podejmują pierwszą próbę zdania części I egzaminu podczas realizacji drugiego roku programu studiów, ale nie jest to wymóg formalny RCP. Regulamin zakłada, że może to nastąpić do 12 miesięcy po zakończeniu nauki na uniwersytecie lub po zakończeniu pierwszego roku programu podstawowego (RCP, 2016d). Zgodnie z regulaminem kandydat może podjąć maksymalnie sześć prób zdania każdej części i ma siedem lat na zdanie całości. Po upływie tego czasu musi powtórnie zdać każdą część. Kandydaci, którzy nie zdołają uzyskać pozytywnego wyniku egzaminu (wykorzystają 6 prób), muszą wybrać inną specjalizację medyczną.

Uzyskanie kwalifikacji MRCP(UK) to proces czasochłonny. Dane dotyczące zdawalności kandydatów w okresie 2003–2011 pokazują, że dyplom MRCP(UK) otrzymuje około 28% podchodzących do egzaminu, a kandydaci potrzebują średnio dwóch prób, by zdać część I, 1,6 próby, aby zaliczyć część II oraz 1,8 próby, by uzyskać wynik pozytywny z PACES (Ludka-Stempień, 2015). Jedynie 8% spośród wszystkich kandydatów przystępujących do egzaminu zdaje wszystkie części za pierwszym podejściem.

Uzyskanie MRCP(UK) wymaga sporych nakładów finansowych ze strony kandydata. Przystąpienie do wszystkich części kosztuje łącznie około 1500 funtów brytyjskich, przy założeniu, że każdą część kandydat zdaje tylko raz, co rzadko się zdarza. Opłata za egzamin pokrywa jedynie koszty egzaminacyjne. Części I i II przeprowadzana jest w centrach egzaminacyjnych; kandydaci odpowiadają na pytania na papierowych arkuszach testowych; wynik jest obliczany komputerowo po zeskanowaniu arkuszy. Egzamin praktyczny odbywa się na oddziałach szpitalnych.

#### 4.2.1. Część I egzaminu

Celem części I jest ocena wiedzy i rozumienia kandydata w zakresie nauk ścisłych, mających odniesienie do medycyny oraz znajomości powszechnych lub ważnych chorób w stopniu umożliwiającym przystąpienie do szkolenia specjalistycznego. Egzamin składa się dwóch trzygodzinnych testów, które odbywają się w jednym dniu. Każdy test zawiera sto pytań, w których tylko jedna z pięciu możliwych odpowiedzi jest prawidłowa. Pytania dotyczą zagadnień z 14 obszarów medycyny ogólnej w proporcjach opisanych w Tabeli 1.

Pytania na teście pochodzą z ogólnej puli. Każde z pytań może zostać użyte co najwyżej pięć razy (nieważne w jakich odstępach czasu), a po piątym razie nie może pojawić się na kolejnym egzaminie przez co najmniej trzy lata.

Komisja egzaminacyjna składa się z 18 wybranych członków college'u. Przed każdą sesją przegląda pytania i analizuje je szczegółowo pod kątem języka, zakresu, trudności, mocy dyskryminacyjnej oraz jakości możliwych do wyboru odpowiedzi. Odrzucane są pytania budzące wątpliwości, zastępuje się je kolejnymi pytaniami z banku pytań. W ten sposób ustalana jest ostateczna zawartość testu, którą zatwierdza cała komisja.

#### Tabela 1. Skład testów egzaminacyjnych części I (łącznie)

Dziedzina	Liczba pytań
Kardiologia	15
Hematologia i onkologia	15
Farmakologia, terapia, toksykologia	16
Nauki ścisłe i medyczne: Biologia komórki Anatomia Metabolizm i biochemia organizmu Fizjologia Genetyka Immunologia Statystyka, epidemiologia i medycyna oparta na dowodach	25 2 3 4 4 3 4 5
Dermatologia	8
Endokrynologia	15
Geriatrya	4
Gastroenterologia	15
Choroby Zakaźne	15
Neurologia	15

Nefrologia	15
Okulistyka	4
Psychiatria	8
Pulmonologia	15
Reumatologia	15
Łącznie	200

Źródło: Ludka-Stempień (2015), za: [www.mrcpuk.org](http://www.mrcpuk.org)

Nie wszystkie odpowiedzi udzielone przez kandydata są uwzględniane w ocenie testu. Około 10% pytań z 200 nie jest ocenianych, ich celem jest zrównywanie statystyczne wyników pomiędzy poszczególnymi sesjami. Pytania uznane za wadliwe również nie są brane pod uwagę przy ostatecznej ocenie.

Wyniki egzaminu są ogłaszane cztery tygodnie od daty jego przeprowadzenia, dotyczy to części I i II. Rysunek 3. przedstawia proces oceny, obejmujący obliczanie wyników oraz wyznaczenie progu punktów wymaganego do zaliczenia egzaminu. Jest to proces wieloetapowy, który zakłada ponowną weryfikację jakości pozycji testowych (tzw. *items*), usunięcie pozycji wadliwych, a pozostawienie tych, które będą punktowane. Wynik po przeliczeniu przedstawiany jest w skali od 0 do 999 punktów.



**Rysunek 3. Proces obliczania wyników części I i II MRCPUK**

Źródło: [www.mrcpuk.org](http://www.mrcpuk.org)

#### 4.2.2. Część II egzaminu

Część II ma formę testu pisemnego, który ma na celu zweryfikowanie, czy kandydat umie zastosować wiedzę medyczną w odniesieniu do konkretnych przypadków pacjentów, choć jeszcze nie w warunkach symulowanych, w których przeprowadzana jest praktyczna część egzaminu (PACES). Pytania dotyczą między innymi analizy wyników pacjenta, stawiania diagnozy i prognozowania. Część ta składa się z trzech pisemnych testów, do których kandydat przystępuje przez dwa kolejne dni. Każdy test liczy 90 pytań w formacie analogicznym do części I. Pytania często zawierają opis przypadku klinicznego, wyniki przeprowadzonych badań, mogą się w nich również znaleźć ilustracje lub zdjęcia. Przykładowe pytania zamieszczono na stronie internetowej egzaminu. Także w przypadku części II komisja egzaminacyjna wybiera pytania, opierając się na danych psychometrycznych ze ściśle określonych dziedzin, w proporcjach odzwierciedlonych w Tabeli 2.

Analogicznie do części I, każde z pytań może zostać użyte maksymalnie pięć razy (nieważne w jakich odstępach czasu), po czym trzeba odczekać trzy lata, zanim zostanie użyte ponownie. W części II egzaminu również zawarte są pytania, które nie są oceniane i służą porównywaniu trudności pomiędzy egzaminami w kolejnych latach.

**Tabela 2. Skład testów egzaminacyjnych części II (łącznie)**



<b>Dziedzina</b>	<b>Liczba pytań ogółem</b>
Kardiologia	25
Dermatologia	13
Endokrynologia i metabolizm	25
Gastroenterologia	25
Geriatryka	10
Hematologia	13
Choroby Zakaźne	25
Neurologia	25
Nefrologia	25
Onkologia i medycyna paliatywna	13
Okulistyka	4
Psychiatria	4
Pulmonologia	25

Reumatologia	13
Terapeutyka i toksykologia	25
Łącznie	270

Źródło: Ludka-Stempień (2015), za: [www.mrcpuk.org](http://www.mrcpuk.org)

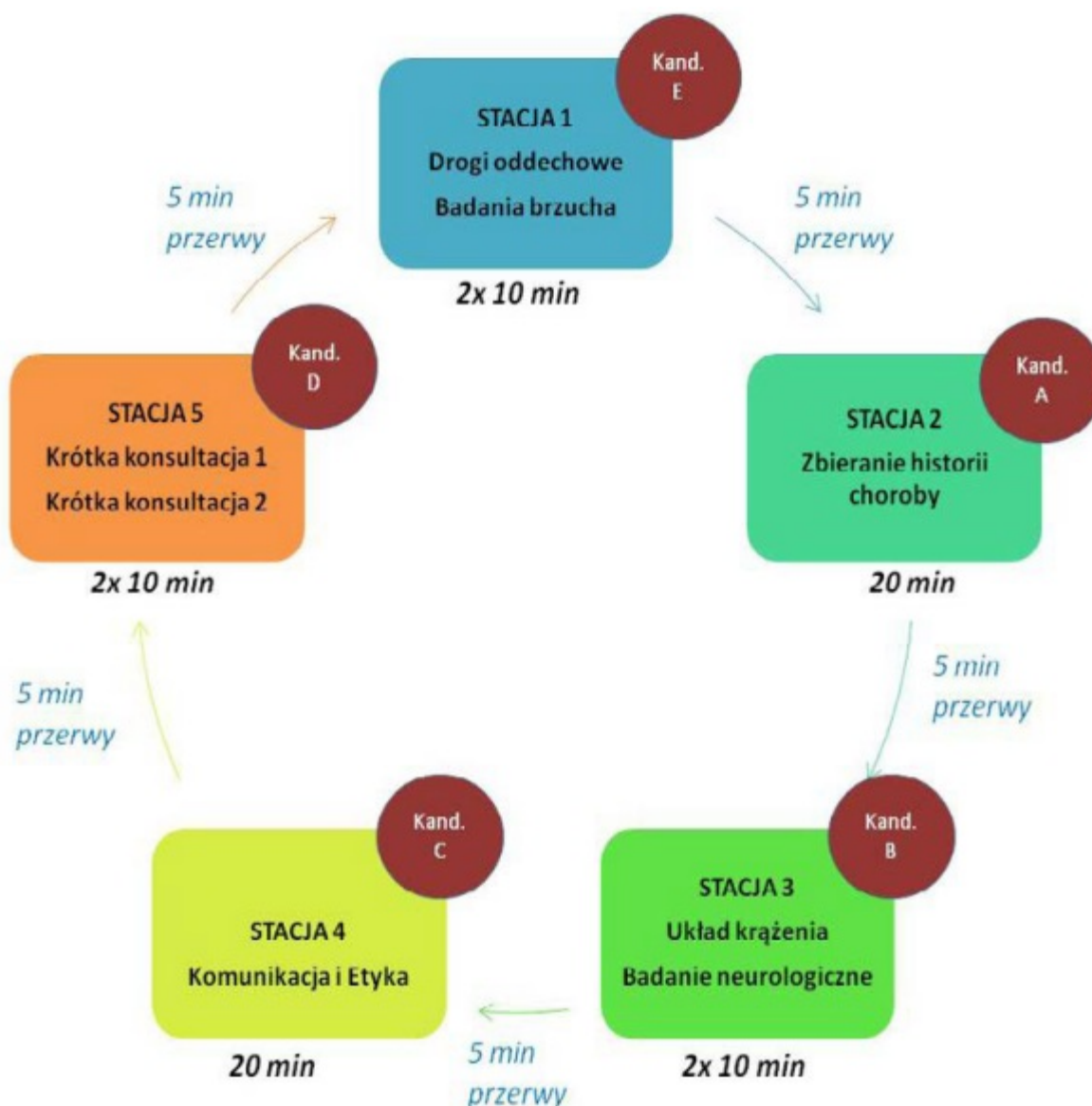
Proces oceniania arkuszy egzaminacyjnych jest analogiczny do stosowanego w przypadku części I. Egzamin odbywają się w wyznaczonych centrach egzaminacyjnych, a obliczanie parametrów poszczególnych pozycji egzaminu, wyników indywidualnych oraz progu gwarantującego zaliczenie egzaminu trwa cztery tygodnie. Po tym czasie kandydaci otrzymują informację o wyniku egzaminu.

#### 4.2.3. Praktyczna ocena umiejętności klinicznych (PACES)

PACES weryfikuje, czy kandydat posiada określone kompetencje niezbędne w praktyce klinicznej, takie jak:

- prowadzenie wywiadu z pacjentem;
- komunikacja wyników pacjentom, ich rodzinom oraz kolegom;
- poprawne przeprowadzenie badania i interpretacja objawów;
- postawienie odpowiedniej diagnozy;
- umiejętność zaplanowania i przedyskutowania natychmiastowych działań i długotrwałych planów leczenia;
- uwzględnienie kwestii etycznych w zgodzie z kryteriami określonymi w *Good medical practice* (GMC, 2013a).

Egzamin trwa obecnie dwie godziny i pięć minut i przeprowadzany jest na pięciu tzw. stanowiskach (*station*). Na każdym stanowisku kandydaci otrzymują zadania, których celem jest ocena ich umiejętności i wiedzy pochodzących z różnych obszarów medycyny wewnętrznej oraz umiejętności komunikacji. W ramach każdego stanowiska opisane są inne przypadki pacjentów z problemami dotyczącymi: (i) układu oddechowego i dolnych partii brzucha (dwa przypadki po 10 minut), (ii) układu nerwowego i krążenia (dwa przypadki po 10 minut), (iii) chorób skóry, zaburzeń równowagi, chorób oczu i układu hormonalnego (cztery przypadki po pięć minut) – łącznie trzy stanowiska. Na dwóch ostatnich stanowiskach kandydaci otrzymują zadanie wywiadu z pacjentem oraz weryfikowane są ich umiejętności komunikacji oraz etyki zawodowej. Na wykonywanie zadania na każdym stanowisku przeznaczono po 20 minut, co odpowiada sytuacji pracy w szpitalu, gdzie tyle czasu lekarz faktycznie otrzymuje na ocenę stanu zdrowia pacjenta. Pomiedzy zadaniami zaplanowano pięć minut przerwy. Koniec i początek każdej sesji jest ogłaszany dzwonkiem. Przed przystąpieniem do zadania na każdym stanowisku kandydat przygotowuje się podczas przerw i na pięć minut przed rozpoczęciem pierwszego zadania. W tym czasie kandydat zapoznaje się także ze scenariuszem zadania, który może zawierać: wyniki pacjenta, wprowadzenie do rozmowy z pacjentem, polecenia. Egzamin jest przeprowadzany jednocześnie dla pięciu kandydatów, z których każdy rozwiązuje zadania przy innym stanowisku. Po wykonaniu zadania, kandydaci przechodzą do innych stanowisk.



**Rysunek 4. Schemat organizacji egzaminu PACES**

Źródło: Ludka-Stempień (2015), za: [www.mrcpuk.org](http://www.mrcpuk.org)

Schemat procedury egzaminu PACES pokazano na Rysunku 4, a jego przebieg można obejrzeć na filmie instruktażowym zamieszczonym przez college na portalu YouTube.

W trakcie egzaminu kandydaci analizują przypadki osób autentycznie chorych lub rolę pacjentów odgrywają zawodowi aktorzy. Pojawiają się oni głównie w zadaniach dotyczących zebrania historii choroby pacjenta oraz związanych z komunikacją i etyką. Istnieją szczegółowe wytyczne określające sposób przeprowadzania egzaminu. Obejmują one również takie elementy jak np. ustawienie mebli w pomieszczeniach egzaminacyjnych (RCP, 2016c).

**Tabela 3. Lista umiejętności ocenianych na egzaminie PACES**

Umiejętność	Badanie brzucha	Układ oddechowy	Zbieranie historii choroby	Układ krążenia	Układ nerwowy	Komunikacja i etyka	Krótką konsultacja 1	Krótką konsultacja 2	Minimalna liczba punktów do uzyskania
A. Badanie fizyczne	X	X		X	X		X	X	14
B. Określenie objawów fizycznych	X	X		X	X		X	X	14
C. Komunikowanie się z pacjentem			X			X	X	X	10
D. Diagnoza różnicowa	X	X	X	X	X		X	X	16
E. Ocena kliniczna	X	X	X	X	X	X	X	X	18
F. Odnoszenie się do obaw pacjenta			X			X	X	X	10
G. Utrzymanie dobrego samopoczucia pacjenta	X	X	X	X	X	X	X	X	28

Źródło: Ludka-Stempień (2015), za: [www.mrcpuk.org](http://www.mrcpuk.org)

Kandydat musi otrzymać minimalną liczbę punktów z każdego zestawu umiejętności, których listę zawiera Tabela 3. Zaznaczono w niej również, w którym zadaniu jest oceniana dana umiejętność. Kandydaci muszą uzyskać łącznie 130 punktów. Jest to wartość wyższa niż suma minimalnych ocen (110 punktów), co oznacza, że zaliczenie egzaminu wymaga osiągnięcia w niektórych zestawach umiejętności liczby punktów powyżej minimum. Przykładowy arkusz ocen, jakim obecnie posługują się egzaminatorzy znajduje się w Załączniku B.

Wykonanie zadań w ramach każdego stanowiska jest oceniane niezależnie przez dwóch egzaminatorów. Jeden z nich rozmawia z kandydatem i zadaje mu pytania. Drugi w tym czasie analizuje przebieg sesji. Egzaminatorzy zamieniają się rolami w trakcie oceny kolejnych kandydatów. Egzaminatorzy nie mogą komentować przebiegu egzaminu w trakcie jego trwania, nie powinni również odpowiadać na pytania kandydata ani zadawać mu pytań naprowadzających. Egzaminatorzy podczas egzaminu w żaden sposób nie mogą wyrażać swojej opinii o kandydacie ani o tym, co mówi lub robi. Procedury mają pomóc w zachowaniu odpowiednich standardów przeprowadzania egzaminu. Egzaminatorzy przechodzą szczegółowe szkolenie oraz otrzymują dokładne pisemne wytyczne (RCP, 2016c).

Skala ocen jest trzystopniowa. Poziom uzyskanej punktacji może być

- satysfakcjonujący,
- na granicy,
- niesatysfakcjonujący.

Egzaminatorzy, którzy zdecydują się wystawić kandydatom ocenę niską lub na granicy, są zobowiązani pisemnie uzasadnić swój wybór, ponieważ uzyskanie tych ocen skutkuje niezaliczeniem egzaminu. Ponadto egzaminatorzy, jeśli uznają to za stosowne, mogą przekazać kandydatowi informację zwrotną.

Egzaminatorem PACES może zostać osoba, która spełnia następujące kryteria:

**» jeśli jest lekarzem z Wielkiej Brytanii:**

- posiada dyplom MRCP(UK);
- jest członkiem honorowym lub kolegią jednego z trzech college'ów medycyny wewnętrznej (członkowie kolegiów muszą posiadać certyfikat ukończenia szkolenia specjalizacyjnego od co najmniej dwóch lat i muszą pracować na stanowisku konsultanta);
- jest zarejestrowana w General Medical Council (GMC);
- uczestniczy w procesie doskonalenia zawodowego, w tym poddaje się rocznej ocenie w miejscu pracy oraz wykazuje dowody na zdobywanie nowych kompetencji.

**» jeśli jest lekarzem spoza Wielkiej Brytanii:**

- jest członkiem honorowym jednego z trzech college'ów medycyny wewnętrznej;
- jest wpisana do rejestru lekarzy oraz posiada prawo do wykonania zawodu (licencji).

**» a ponadto (kryteria wspólne dla wszystkich lekarzy):**

- praktykowała medycynę w ciągu ostatnich dwóch lat w szpitalu lub innych placówkach medycznych;
- na co dzień ocenia, bada i zajmuje się przypadkami klinicznymi ze spektrum zagadnień, które obejmuje PACES;
- nadzoruje szkolenie lekarzy na poziomie głównego szkolenia specjalistycznego (CMT) lub analogicznego, którzy zajmują się pacjentami o schorzeniach ze spektrum zagadnień obejmowanych przez PACES;
- odbyła szkolenie z równego traktowania oraz różnic kulturowych w ciągu ostatnich trzech lat;
- jest gotowa do egzaminowania przynajmniej 30 kandydatów rocznie (sześć cykli egzaminacyjnych, dwa dni).

Egzaminatorzy przechodzą specjalne szkolenie mające na celu przygotowanie do oceniania z zachowaniem spójności wystawianych ocen oraz utrzymanie rzetelności egzaminu. Szkolenia te odbywają się podczas sesji egzaminacyjnych PACES. Kandydaci na egzaminatorów najpierw uczestniczą w spotkaniu

przedegzaminacyjnym, w trakcie którego faktyczni egzaminatorzy ustalają, jakie fizyczne objawy powinien zbadać egzaminowany, jakie pytania mogą paść ze strony egzaminatorów i jak będą oceniać przebieg egzaminu. Następnie kandydat na egzaminatora dołącza do jednego ze stanowisk jako obserwator i podobnie jak faktyczni egzaminatorzy ocenia umiejętności zdającego. Arkusz wypełniony przez kandydata na egzaminatora jest później porównywany z arkuszami egzaminatorów i przez nich omawiany. Po trzech cyklach egzaminacyjnych (czyli na koniec dnia) ocenia się arkusze kandydata na egzaminatora oraz przekazuje się mu informację zwrotną od pozostałych egzaminatorów. Jeśli kandydat na egzaminatora podtrzymuje chęć bycia egzaminatorem i został dobrze oceniony, wówczas dostaje możliwość samodzielnego oceniania PACES (RCP, 2016b).

#### 4.2.4. Rzetelność i trafność wyników egzaminu

Wskaźniki rzetelności dla wszystkich części egzaminu utrzymują się na wysokim poziomie tj. powyżej 0,90 dla części I oraz II i 0,80 dla PACES. Jego trafność potwierdza się przede wszystkim badaniem istotności związków statystycznych pomiędzy wynikami poszczególnych części MRCP(UK) a wynikami 17 innych egzaminów pisemnych i dwóch egzaminów klinicznych z zakresu innych specjalności lekarskich, a także związku z ocenami rocznymi oraz przypadkami odebrania licencji lub ograniczenia prawa do wykonywania zawodu przez GMC. Związki te były silniejsze w przypadku egzaminów pisemnych oraz słabsze w przypadku miar będących praktyczną oceną umiejętności klinicznych (Ludka-Stempień, 2015). Kluczowym wnioskiem z tych analiz jest to, że osoby które otrzymują niskie oceny na egzaminie, mają większe szanse na późniejsze problemy z licencją, mniejsze szanse na dobre oceny w trakcie szkolenia oraz mniejsze szanse na dobre wyniki egzaminów kończących szkolenie, co potwierdza trafność prognostyczną egzaminu.

## Doskonalenie procesu weryfikacji – kwalifikacja medyczna MRCP(UK)

### DZIŚ

**Weryfikacja trafności i rzetelności**  
Przez ostatnie 30 lat nieustannie trwa proces weryfikacji trafności i rzetelności egzaminu MRCP(UK).



### 2001

**Wprowadzenie PACES**  
Częścią egzaminu jest praktyczna ocena umiejętności klinicznych (PACES). Weryfikuje ona również umiejętności komunikacyjne.



### 1969

**Wspólny egzamin**  
MRCP - jako spełniający wymogi wszystkich college'y - stał się wspólnym egzaminem, w którym długie pytania otwarte zastąpiono krótkimi, podzielonymi na grupy. Położono nacisk na znajomość powszechnych przypadków chorobowych zamiast rzadkich i skomplikowanych.



### 1966

**Test odsiewający**  
RCP wprowadza test odsiewający, który pozwala na znaczne zwiększenie wydajności przeprowadzania egzaminów oraz na weryfikację wiedzy kandydatów przed kluczową częścią praktyczną.



### 1885

**Nowy egzamin w Glasgow**  
Ze względu na zarzuty o niespełnianiu standardów jakości College w Glasgow wprowadza nowy egzamin bazujący na doświadczeniach college'u z Edynburga.



### 1881

**Pierwsze egzaminy**  
College w Edynburgu przeprowadza pierwszy egzamin, składający się z części pisemnej i części praktycznej trwających łącznie 3 dni.



### 1858

**Licencje nadawane centralnie**  
W życie wchodzi ustawa, na mocy której nadawaniem licencji zarządza General Medical Council. College'e zajmują się nauczaniem i egzaminowaniem.



### 1518

**Powstanie college'y lekarskich**  
Powstaje The Royal College of Physicians of London – pierwszy college zrzeszający lekarzy z wstępnym egzaminem ustnym (RCP). Zasady przystąpienia do RCP były niejasne i nie egalitarne. W kolejnych latach powstają podobne college'e w Edynburgu i Glasgow.



### 4.3. Wieki doświadczenia prowadzące do MRCP(UK)

Obecna formuła egzaminu jest wynikiem wielu lat tradycji przeprowadzania egzaminów lekarskich z założeniem ich stałego udoskonalania. Wprowadzane zmiany były wynikiem zarówno rozwoju wiedzy i postępu technologicznego, jak i zachodzących przemian społecznych. Opis tych zmian wyjaśnia, dlaczego certyfikat MRCP(UK) jest uważany za prestiżowy.

#### 4.3.1. The Royal College of Physicians of London

Powstanie college'ów jako instytucji formalnie nadających kwalifikację lekarza stanowiło pierwszą próbę zapewnienia wysokiej jakości usług medycznych.

The Royal College of Physicians of London był pierwszym college'em zrzeszającym lekarzy. Został założony w 1518 roku przez króla Henryka VIII na mocy edyktu, który został wydany w odpowiedzi na petycję złożoną przez siedmiu medyków świadczących usługi w rejonie ówczesnego Londynu. W tamtych czasach usługi lekarzy zarezerwowane były jedynie dla arystokracji. Pozostała część społeczeństwa korzystała z wszelkiej pomocy zielarzy, balwierzy, chirurgów, często zwykłych szarlatanów. Wspomniana petycja do króla miała dwa cele: zapewnienie jakości świadczonych usług poprzez wyeliminowanie osób nieposiadających odpowiednich kompetencji, ale też ograniczenie konkurencji na rynku usług medycznych. Akt królewski zapewniał college'owi wyłączność na świadczenie usług medycznych w obrębie miasta Londynu.

Wkrótce po ustanowieniu college'u w Londynie, w 1599 roku, powstał The Royal College of Physicians and Surgeons of Glasgow. Znacznie później, bo w 1681 roku założono The Royal College of Edinburgh. Do 1968 roku college'e te funkcjonowały całkowicie niezależnie, miały swoje strefy wpływów oraz, co istotne, oferowały indywidualne członkostwo. Oznaczało to również, że miały własne egzaminy, które w pewnym stopniu były do siebie podobne.

Egzaminy wstępne do college'u londyńskiego mogą służyć za przykład typowych ówczesnych egzaminów lekarskich. College w Londynie oferował dwa rodzaje członkostwa i dwa tytuły: licencjata (*licenciate*) oraz członka (*fellow*). Obie grupy różniły się znacząco w uprawnieniach – licencjaci mieli prawo do praktykowania medycyny, podczas gdy pełnoprawni członkowie mieli dodatkowo wpływ na ustanawianie przepisów, do których musieli się stosować wszyscy członkowie college'u, oraz mieli prawo do zasiadania w jego władzach oraz głosowania w wyborach tych władz (Waddington, 1973).

Do college'u przyjmowano z polecenia, po przedstawieniu dyplomów z renomowanych uczelni lub na podstawie dowodów prowadzenia dobrej praktyki poza rejonem college'u, np. w innym kraju. Egzamin na pełnoprawnego członka przeprowadzano ustnie, a ocenę wystawiali tzw. cenzorzy (których można uznać za odpowiedników egzaminatorów). Po tym następowało głosowanie. Nie zachowały się zapiski ze szczegółowymi pytaniami, które były zadawane egzaminowanemu. Zasady przystąpienia do college'u można zatem uznać za co najmniej niejasne, a na pewno nie egalitarne. Wiadomo, że pełne członkostwo wymagało co najmniej czterech lat praktyki medycznej. By to umożliwić, egzamin można było zdać warunkowo przed odbyciem koniecznej praktyki, a następnie po czterech latach egzamin musiał zostać powtórzony. Takich członków warunkowych nazywano kandydatami.

Z oczywistych względów college miał niewielu członków. W XVI wieku najwyższa zanotowana ich liczba wynosiła 31 (Clark, 1965). W 1708 roku zaledwie 57 osób miało pełne prawa członkowskie. Dodatkowo było siedmiu kandydatów i trzech członków honorowych. Licencjatorów, dla porównania, było 39. Liczby te były jeszcze niższe w 1746 roku (RCP i Cuthbertson, 2008). By zmienić proporcje w relacji liczby członków do licencjatorów, z powodu niezadowolenia tych ostatnich z ograniczenia ich praw, w 1771 roku wprowadzono nowy statut, zgodnie z którym, jeśli licencjat praktykował przynajmniej przez siedem lat i miał powyżej 35 lat, to mógł być nominowany na kandydata przez dowolnego członka college'u. Zgodnie z zapisami musiał on wówczas zdać egzamin z medycyny (niestety, nie wyszczególniono jego zakresu) oraz ze znajomości pism filozofów takich jak Hipokrates, Galen czy Areteusz. Nadal jednak uznawano, że kandydat zdał egzamin, jeśli członkowie college'u potwierdzili to w głosowaniu. Tę formułę utrzymano do połowy XIX wieku.



#### 4.3.2. Ustawa z 1858 roku i jej konsekwencje

Powołanie General Medical Council jako instytucji przyznającej licencje lekarskie niezależnie od college'ów oraz próba standaryzacji egzaminów były kolejnymi krokami na drodze do zapewnienia jakości kwalifikacji medycznych.

Na początku XIX wieku Londyn liczył już około pół miliona mieszkańców. Dotychczasowa liczba lekarzy nie spełniała potrzeb tak dużej populacji i zarówno ta sytuacja, jak i interesy samych członków college'u wywierały na londyński college presję, by zwiększyć liczbę licencjonowanych medyków (Waddington, 1973). Czasy wiktoriańskie to także okres rozwoju wiedzy medycznej, rozpowszechniania się idei profesjonalizacji, ale też wzrastającej troski o zdrowie publiczne. W tym kontekście duże znaczenie miało uchwalenie ustawy o usługach medycznych (*Medical act*) w 1858 roku (Roberts, 2009).

Ustawa ta została przyjęta przez środowisko z umiarkowanym entuzjazmem. Do tej pory college'e sprawowały absolutną władzę nad procesem nadawania licencji, ponieważ jej uzyskanie było równoznaczne z członkostwem w college'u. Z chwilą przyjęcia ustawy przez parlament kompetencje college'ów w tym zakresie zostały ograniczone, przez co utraciły one swoje strefy wpływów. Procesem nadawania licencji od tej chwili zarządzać miały nie college'e, tylko nowa instytucja: General Medical Council (GMC), zaś college'e, mimo że utrzymały swoją niezależność, stały się instytucjami nauczającymi i egzaminującymi. Można było zdać egzamin w Glasgow, otrzymać licencję GMC i praktykować w Londynie lub gdziekolwiek indziej na terenie Wielkiej Brytanii.

Ze względu na wymienioną ustawę formę zmienił też egzamin członkowski. Wprowadzono tytuł członka zwykłego (*member*) oraz honorowego (*fellow*). Egzamin wstępny do college'u (RCP) był egzaminem na członka zwykłego (stąd akronim MRCP), zaś członkiem honorowym zostawało się w uznaniu zasług dla środowiska medycznego. Członków honorowych wybierano spośród członków zwyczajnych w głosowaniu. Nowy egzamin przeprowadzony został po raz pierwszy w 1859 roku. W 1867 roku przywrócono tytuł licencjata, który został zarezerwowany dla studentów medycyny (Cooke, 1972).

Egzamin ten został bardziej sformalizowany i w tej postaci stał się podstawą obecnego MRCP(UK). W tamtym czasie składał się on z czterech części: pisemnych esejów z teoretycznej wiedzy z fizjologii, patologii (każdy z nich z tłumaczeniem z łaciny i greki) oraz z egzaminu ustnego z wykorzystania i praktyki medycyny (RCP i Cuthbertson, 2008). Egzamin odbywał się etapami co trzy miesiące. Zwolnienie z egzaminu pisemnego było możliwe tylko dla tych kandydatów, którzy zostali uznani przez college za niezwykle zdolnych, jednak nie zostało dokładnie określone, co konkretnie miało to oznaczać. W 1867 roku wprowadzono zmianę, która była wynikiem rekomendacji GMC, by egzamin oceniała para egzaminatorów, a nie jeden jak dotychczas (Cooke, 1972).

Wydania „British Medical Journal” z lat 1879 oraz 1880 (GMC, 1879; 1880) pokazują, jakie były wymagania w stosunku do kandydatów do college'u, jednak nie znajdują się w nich informacje na temat procedur oceniania na egzaminie. Można uznać, że zapewnianie jakości egzaminu leżało wówczas po stronie egzaminatorów. Osoby kandydujące na członka college'u musiały mieć ukończone 25 lat, studiować nie krócej niż pięć lat, z czego cztery na renomowanej uczelni uznanej przez college. Wiadomo też, że w tamtym okresie egzamin składał się z trzech części, oraz że pojawiły się pytania z nowych obszarów medycyny.

Część I obejmowała pytania pisemne z anatomii i fizjologii oraz egzamin ustny z patologii i przygotowania preparatów medycznych. Ta część trwała łącznie dwa dni. Część II trwała cztery dni, podczas których odbywał się egzamin pisemny i ustny z medycyny, chemii w patologii, farmacji i toksykologii; pisemny egzamin z chorób kobiecych i położnictwa, pisemny egzamin z anatomii i chirurgii, praktyczny egzamin na oddziale szpitalnym oraz ustny egzamin z chirurgii i położnictwa. Trzeci egzamin trwał łącznie pięć dni i składał się z kolejnej serii pytań pisemnych i egzaminu praktycznego przeprowadzanego na oddziale szpitalnym.

W 1893 roku została opublikowana pierwsza wzmianka dotycząca procedury oceniania kandydatów w trakcie egzaminu. Doktor William Dickinson, emerytowany cenzor (czyli niejako egzaminator) skrytykował zmienne standardy oceniania kandydatów oraz nazbyt łagodne oceny i zasugerował zunifikowany sposób oceniania odpowiedzi kandydatów (Cooke, 1972). Na skutek tej krytyki w czerwcu 1894 roku college wprowadził nowy,

punktowy system ocen: maksimum 50 punktów za każdy z esejów, 20 punktów za znajomość języka oraz 120 punktów za egzaminy ustne.

W tym czasie college'e Edynburga i Glasgow wprowadziły swoje własne egzaminy. Egzamin college'u edynburskiego pierwszy raz został przeprowadzony w 1881 roku i składał się z pisemnego zadania z zakresu praktyki medycyny i wybranej specjalizacji medycznej (Fleming i in., 1974) oraz egzaminu praktycznego, który składał się z tzw. długich przypadków (*long cases*) i krótkich przypadków (*short cases*), które odnosiły się do rodzaju przypadłości pacjenta: chronicznej lub ostrej. Łączny czas trwania egzaminu wynosił trzy dni.

College z Glasgow od początku istnienia miał własny system egzaminacyjny, jednak po wejściu w życie ustawy z 1858 roku nie organizował swoich egzaminów i przez 27 lat przyjmował każdego, kto mógł wykazać się tytułem renomowanej brytyjskiej lub zagranicznej uczelni, pod warunkiem, że ten sam kandydat otrzymał w głosowaniu dwie trzecie głosów członków honorowych. Po 1885 roku ze względu na zarzuty o niespełnianie standardów jakości w Glasgow wprowadzono nowy egzamin oparty na doświadczeniach college'u z Edynburga. Dyskusje pomiędzy członkami college'u londyńskiego ze szkockimi na temat jakości egzaminów w Szkocji odbywały się publicznie, na łamach „British Medical Journal” i do dziś stanowią interesującą lekturę.

#### 4.3.3. Początki XX wieku

Nowe osiągnięcia technologiczne w medycynie oraz zmiana formuły egzaminu przed wybuchem II wojny światowej stanowiły wstęp do znacznie poważniejszych przemian.

Początek XX wieku przyniósł nowe osiągnięcia technologiczne oraz szybki rozwój wiedzy medycznej: odkryto promieniowanie rentgenowskie, hormony, powstały pierwsze szczepionki. Początkowo nie wpłynęło to na zmiany w formule egzaminu MRCP, chociaż dyskutowano nad wprowadzeniem nowych zakresów wiedzy oraz zmian w regulaminie egzaminu. W latach 1916–1924 stopniowo zmieniał się sposób egzaminowania z języków, a ostatecznie w 1925 roku wydzielono osobny egzamin. Weryfikowanie umiejętności językowych porzucono całkowicie w latach sześćdziesiątych XX wieku.

Na sposób egzaminowania znacząco wpłynęły doświadczenia z czasu II wojny światowej. W okresie międzywojennym zdawalność egzaminu MRCP była niska. W 1933 roku spośród 75 kandydatów zdało jedynie ośmiu. W 1936 roku sukces odniosło 25 ze 130 kandydatów. Zgodnie ze statystykami zdawalność w tym okresie wynosiła ok. 25% (Cooke, 1972). Tuż przed wybuchem II wojny światowej, w 1937 roku, wprowadzono nową formułę egzaminu. MRCP składał się wówczas z dwóch pisemnych części (po cztery otwarte pytania w każdej) oraz z dwóch części ustnych. Nie opublikowano przykładowych arkuszy, zatem trudno ocenić, czym różniły się te pytania od poprzednich, jednak egzamin trwał krócej.

W czasie II wojny światowej w związku z wzrastającym zapotrzebowaniem na lekarzy oraz zwiększeniem się dostępu do edukacji wyższej liczba chętnych do wstąpienia do college'u znacząco wzrosła, co jednak ujawniło niedobory w gronie/kadrze egzaminatorów. Zwiększenie ich liczby nastąpiło dopiero w 1949 roku.

#### 4.3.4. Lata sześćdziesiąte XX wieku i wzrost naboru na studia wyższe

W okresie powojennym wprowadzono do MRCP test odsiewający, pozwalający na znaczne zwiększenie wydajności przeprowadzania egzaminów oraz na weryfikację wiedzy kandydatów przed kluczową częścią praktyczną. Niedługo potem wprowadzono jeden wspólny egzamin pod nazwą MRCP(UK), w którym długie pytania otwarte zastąpiono krótkimi, podzielonymi na grupy. Położono nacisk na znajomość powszechnych przypadków chorobowych zamiast rzadkich i skomplikowanych.

Do roku 1963 egzamin nadal składał się z kilku części pisemnych i ustnych, tak jak zostało to ustanowione w 1858 roku, a jedynie części pisemne stały się mniej wymagające. Panowała wówczas opinia, że egzamin sprawdza raczej znajomość chorób rzadkich niż powszechnych. Kolejnym problemem była zdawalność, która była niższa niż aktualne zapotrzebowanie społeczeństwa na usługi medyczne. Mała liczba przyjmowanych kandydatów na szkolenia specjalistyczne wynikała wprost z formuły egzaminu. Każdy, kto spełniał kryteria

formalne, podchodził do wszystkich części MRCP, co powodowało, że liczba kandydatów do testów pisemnych i praktycznych była identyczna, a to zmniejszało wydajność całego systemu i efektywność egzaminowania.

Ponieważ część praktyczną egzaminu zawsze uznawano za kluczową, wiedzę kandydatów należało zweryfikować jeszcze przed tym etapem. Racjonalną decyzją było podzielenie MRCP na dwie części, przy czym tylko druga składała się z testu pisemnego i egzaminu ustnego. Część I tego egzaminu stała się testem odsiewającym. W nowej formie składała się już nie z długich pytań otwartych, tylko z krótszych z dwiema możliwymi odpowiedziami: „tak” lub „nie”. Do oceniania w tym teście wykorzystano nową zdobycz techniki, czyli komputer Uniwersytetu w Newcastle. Pytania egzaminacyjne były nadal układane przez egzaminatorów college'u. W 1966 roku powstał bank pytań i utworzono w RCP specjalną grupę zajmującą się tworzeniem nowych pytań. Według dostępnej literatury wiedzę psychometryczną po raz pierwszy wykorzystano dopiero w 1967 roku. Oceniono wówczas moc dyskryminacyjną oraz wartość predykcyjną egzaminu.

W 1969 roku część II również została zmodyfikowana. Zamiast długich pytań otwartych, które uznano za najsłabszą i bardzo czasochłonną formę egzaminowania, wprowadzono krótkie pytania otwarte uszeregowane w trzy grupy: historia przypadków, interpretacja danych oraz materiały graficzne. Odpowiedzi na te pytania osobiście oceniali egzaminatorzy (RCP i Cuthbertson, 2008), ale klucz był wcześniej uzgadniany przez osoby oceniające prace, zaś punktację ustalano na podstawie prawdopodobieństwa wystąpienia możliwych odpowiedzi. Przykład takiego pytania można znaleźć w publikacji Fleminga (Fleming i in., 1974).

Egzamin ustny, czyli praktyczny w części II, nadal obowiązywał, ponieważ dawał on egzaminatorom możliwość bezpośredniej weryfikacji wiedzy i umiejętności kandydata. Zmieniła się jednak jego forma. Tradycyjnie kandydaci mieli do czynienia z jednym, długim przypadkiem medycznym oraz kilkoma krótkimi. Ponieważ jest niewiele chorób, które charakteryzują się skomplikowanym przebiegiem, długie przypadki na egzaminie były wśród kandydatów powszechnie znane. Z tego powodu trudno było go uznać za rzetelny, bowiem stawiana przez kandydatów diagnoza lekarska podczas egzaminu nie dawała wiedzy o faktycznych umiejętnościach lekarskich. Długie przypadki zostały zatem całkowicie wyeliminowane z pytań egzaminacyjnych. Ponadto, jak już wcześniej wspomniano, MRCP był uznawany za egzamin z chorób rzadkich lub przypadków nietypowych. W nowej koncepcji egzamin obejmował przypadki chorób powszechnie występujących, ale czasem dających nietypowe objawy. Zmiana ta miała poszerzyć pulę przypadków egzaminacyjnych. MRCP w nowej formule spełniał wymogi trzech college'ów medycyny wewnętrznej, czyli Londynu, Glasgow i Edynburga. W 1969 roku MRCP stał się egzaminem wspólnym dla wszystkich i zmienił nazwę na MRCP(UK).

W latach 1950–1970 do egzaminu MRCP dołożono część III dla niektórych specjalizacji, np. psychiatrii czy pediatrii, którą jednak zlikwidowano w momencie, gdy powstały osobne college'e prowadzące te specjalizacje.

#### 4.3.5. Od lat siedemdziesiątych XX wieku do dnia dzisiejszego

Zmiany technologiczne lat siedemdziesiątych ułatwiły nie tylko sprawdzanie, ale również ocenę psychometryczną egzaminu, co zaowocowało długotrwałą i szczegółową weryfikacją trafności i rzetelności MRCP(UK). Początek XXI wieku przyniósł znaczącą zmianę w postaci wprowadzenia praktycznej oceny umiejętności klinicznych (PACES), którą wykorzystano również do weryfikacji coraz bardziej docenianych w zawodzie lekarza umiejętności komunikacyjnych.

Częstotliwość wprowadzania zmian do formuły egzaminu w tym okresie była duża, podobnie jak ich zakres. Wynikało z postępu w naukach medycznych, technologii oraz z zachodzących przemian społecznych i politycznych. Dotychczas egzaminy były oceniane osobiście przez egzaminatorów. Pomimo ustalonych kryteriów i dbania o spójność ocen, był to proces czasochłonny i wymagający zaangażowania wielu osób. Ponadto obliczanie statystyk i parametrów służących do oceny jakości MRCP(UK) wymagało wielu wysiłków. Dostępność komputerów sprawiła, że sprawniej szło nie tylko samo podliczanie wyników egzaminu, ale również o wiele łatwiej można było przeprowadzać analizy psychometryczne. Lata 1984–2015 to także okres ciągłej weryfikacji trafności egzaminu MRCP(UK) w celu utrzymania jego wiarygodności i jakości

##### 4.3.5.1. Format pytań i punktacja

Wyniki analiz egzaminów testowych wskazywały, że pytania zamknięte w formacie „tak/nie” ułatwiają kandydatom przypadkowe udzielanie odpowiedzi. Z tego powodu na początku lat siedemdziesiątych arkusze egzaminacyjne zostały zmodyfikowane, a do systemu oceniania wprowadzono ujemne punkty w celu zniechęcenia uczestników egzaminu do zgadywania.

W 1984 roku ustalono, że część I będzie składać się 60 pytań z pięcioma podpunktami w formacie „tak/nie”, co łącznie dawało 300 ocenianych pozycji testu, co obowiązywało do 2002 roku.

Część I powstała w celu preselekcji kandydatów, zatem część II mogła odnieść się już do bardziej szczegółowych zagadnień i stała się egzaminem ukierunkowanym na sprawdzenie praktycznej wiedzy klinicznej. Na wyniki tej części składały się łącznie punkty z testu oraz punkty z egzaminu klinicznego. Pytania testowe (analogiczne do części I) zaczęły dotyczyć takich aspektów, jak: ocena sytuacji pacjenta, umiejętność czytania wyników, diagnozy, podejmowania decyzji oraz prognozy dla pacjenta. Część kliniczna pozostała bez zmian w stosunku do lat poprzednich.

#### 4.3.5.2. Ocena psychometryczna

Ważnym elementem zapewniania jakości egzaminu jest ocena psychometryczna. W latach 1984–2001 rzetelność egzaminu mierzona współczynnikiem rzetelności (alfa Cronbacha) wynosiła ok. 0,87 (McManus i in., 2003) i stopniowo wzrastała, by osiągnąć poziom 0,92. Rzetelność części II była niższa i w latach 2002–2005 współczynniki rzetelności wynosiły 0,73 do 0,83 (Tighe, McManus, Dewhurst, Chis i Mucklow, 2010) Standard Error of Measurement (SEM). Do 2005 roku część II składała się z dwóch testów po 100 pytań, ale ze względu na niską rzetelność liczba pytań wzrosła do 270, co jest zgodne z formułą Spearmana-Browna (por. Hornowska, 2005). Ostatecznie rzetelność pomiaru części II stała się zadowalająca i osiągnęła poziom 0,90.

W ramach oceny jakości egzaminu badano także trafność, przy czym weryfikowano nie tylko hipotezy związane z zagadnieniami społecznymi, jak na przykład ewentualną dyskryminację ze względu na płeć (Dewhurst i in., 2007; McManus i Elder i in., 2008), lecz także korelację pomiędzy wynikami egzaminu MRCP(UK) a efektami pracy lekarzy lub wynikami dalszych szkoleń (tzw. trafność prognostyczna; Levy i in., 2011a; Ludka-Stempień, 2015). Celem tych analiz było zapewnienie pełnej transparentności instytucji i egzaminu MRCP(UK) oraz zapewnienie jego wiarygodności.

Niektóre z przeprowadzonych analiz miały znaczący wpływ, jeśli nie na samą formułę egzaminu, to na procedurę egzaminacyjną. Przykładowo, projekt zrealizowany przez zespół McManusa (McManus i in., 2005) analizujący umiejętności poszczególnych roczników przystępujących do egzaminu w latach 1985–2001 wykazał, że absolwenci uczelni z lat 1985–1996 mieli znacznie wyższe kompetencje niż absolwenci kolejnych roczników. Przyczyny tego zjawiska nie zostały jednoznacznie sprecyzowane w artykule, ponieważ potencjalnie wiele czynników mogło mieć wpływ na taki stan rzeczy. Jednak wyniki uzasadniały użycie przez RCP nowych metod ustalania progu punktów pozwalających zdać egzamin w sposób inny niż opierający się na odniesieniu do populacji (tzw. *norm-referencing*).

Początkowo ustalono, że za minimalny próg punktów dla testów (części I i II) uznany zostanie stały procent poprawnych odpowiedzi. Powodowało to jednak, że kandydaci kwestionowali wiarygodność egzaminu, bo system oceny był całkowicie oderwany zarówno od poziomu trudności samego egzaminu, jak i od poziomu umiejętności w populacji, którą miał badać. Ponieważ pytania były zmieniane w kolejnych edycjach testu, ten sposób ustalania progu nie pozwalał przewidzieć, ilu kandydatów zda egzamin. Alternatywnym podejściem było ustalenie nie tyle minimalnej liczby punktów uzyskanej na egzaminie, ile maksymalnej liczby kandydatów, która może zdać egzamin. Limit ten ustalono na poziomie 35% wszystkich podejmujących próbę (*norm-referencing*). Jednak rozwiązanie to nadal nie gwarantowało, że kandydaci przyjmowani do college'ów będą mieli wystarczający poziom wiedzy, a jak pokazywały wspomniane w poprzednim akapicie badania McManusa, poziom wiedzy i umiejętności kandydatów kończących studia medyczne spadał.

W związku z powyższym wprowadzono ustalanie progu zdawalności metodą Angoffa, a następnie metodą Hofstee. Po dokładny opis obu metod warto sięgnąć do publikacji Norciniego (Norcini, 2003). Metoda Angoffa polega na tym, że grupa kompetentnych egzaminatorów ocenia to, jakie umiejętności i wiedzę powinien posiadać kandydat, by otrzymać graniczną liczbę punktów, umożliwiającą zaliczenie egzaminu (*borderline*

*candidate*) oraz – biorąc pod uwagę pytania – jakie jest prawdopodobieństwo, że taki kandydat odpowie na każde z pytań. Na tej podstawie wylicza się średnią z ocen członków komisji, która daje punktację graniczną.

Metoda Hofstee z kolei polega na tym, że nie ocenia się wszystkich pozycji testu, a jedynie wskazuje zakres granicznej punktacji pozwalającej zdać egzamin oraz procent pożądaną liczbę kandydatów, jaka docelowo ma uzyskać pozytywny wynik (Hofstee, Berge i Hendriks, 1998; Norcini, 2003). Metoda Hofstee funkcjonowała do końca 2008 roku dla części I oraz do końca 2009 roku dla części II.

Ostatecznie wprowadzono najlepszą i najnowszą metodę ustalania punktacji granicznej, czyli tzw. zrównywanie statystyczne (*statistical equating*). Metoda ta jest dość skomplikowana, ale została dobrze opisana przez Skaggs i Lissitz (1986) oraz Mosesa i współpracowników (Moses, Deng i Zhang, 2010) *The Use of Two Anchors in Nonequivalent Groups With Anchor Test NEAT*. Opiera się ona na parametrach poszczególnych pozycji i pozwala uwzględnić trudność każdego z pytań oraz możliwości grupy kandydatów.

Kolejne badanie psychometryczne dotyczące wielokrotnych prób zdania egzaminu wykazało, że istnieje grupa kandydatów, w przypadku której umożliwienie podejmowania wielokrotnych prób zdania egzaminu, może prowadzić do zaliczenia go dzięki „szczęściu” i otrzymania kwalifikacji pomimo braku faktycznych kompetencji (McManus i Ludka, 2013). Skutkiem tego badania było ograniczenie w 2014 roku maksymalnej liczby prób podejścia do poszczególnych części egzaminu do sześciu.

#### 4.3.5.3. Wyodrębnienie PACES oraz badanie jego trafności i rzetelności

W 2001 roku część II została oficjalnie podzielona na egzamin pisemny oraz na tzw. PACES (czyli praktyczną ocenę umiejętności klinicznych. PACES stał się odrębną częścią MRCP(UK). Do niedawna można było podejść do tej części egzaminu tylko po zdaniu części II.

Wprowadzenie PACES było wynikiem szerszych zmian w myśleniu i publicznej dyskusji o egzaminowaniu faktycznych umiejętności. Zarzucano MRCP(UK), że formuła egzaminu i prawdopodobieństwo jego zdania są podobne do zasad gry hazardowej (London, Parry, Clesham i O'Brien, 1994). Celem nowej formuły egzaminu klinicznego było osadzenie go w realiach współczesnej medycyny i myśleniu opartym na dowodach naukowych oraz dostosowaniu go do współczesnych metod weryfikacji umiejętności, funkcjonujących już na uczelniach (Dacre, Besser i White, 2003). PACES naśladował zatem proces OSCE (*objective structured clinical examination*), czyli ustrukturyzowany obiektywny egzamin kliniczny, który w latach dziewięćdziesiątych XX wieku stał się popularną metodą oceny umiejętności studentów na uczelniach medycznych. OSCE to seria krótkich badań i zadań klinicznych, które powstały po to, by szybciej i sprawniej oceniać umiejętności studentów medycyny (Harden, Stevenson, Downie i Wilson, 1975). Co do zakresu ocenianych umiejętności klinicznych, to niewiele się zmieniło, z tym, że większy nacisk położono na umiejętności komunikacji i gromadzenia historii choroby, co było też odzwierciedleniem zmian wprowadzonych do *Good medical practice*.

PACES od początku miał zbliżoną formułę do obecnej. W 2009 roku wprowadzono zmiany, które polegały przede wszystkim na odejściu od dotychczasowego sposobu oceniania umiejętności specyficznych na rzecz kompetencji ogólnych, na przykład: umiejętności komunikacji z pacjentem, radzenia sobie z jego obawami, przeprowadzenia badania fizycznego. Kandydat musi otrzymać minimalną liczbę punktów na każdej ze skal określających kompetencje, a nie, jak dotychczas, uzyskać dobry wynik w trzech zadaniach z pięciu, co pozwalało kompensować niedostatki kompetencji w komunikacji umiejętnością przeprowadzenia dobrego badania fizycznego lub na odwrót (Elder i in., 2011).

Ze względu na praktyczną formę egzaminu weryfikacja jego jakości była zadaniem znacznie trudniejszym niż w przypadku testów. Niemniej także dla tej części potwierdzono wiarygodność jego wyników. Współczynnik rzetelności testu został obliczony przez McManusa (McManus, Thompson i Mollon, 2006) i wynosił 0,82, co jest wynikiem zadowalającym, choć nie doskonałym. Wyniki przedstawione przez Dacre i innych (2003) wskazywały na wysoką rzetelność ocen stawianych przez egzaminatorów, co udowodniało wysoki poziom standaryzacji procedury egzaminu. Przytaczając wyniki tych badań, zaledwie w przypadku 2,2% egzaminatorów pracujących w parach przyznawana ocena różniła się o więcej niż dwa punkty, a 60,7% egzaminatorów była w całkowitej zgodzie co do oceny (Dacre i in., 2003). Badania McManusa z 2006 roku

pokazywały ponadto, że 87% wariacji wyników egzaminacyjnych było spowodowane faktycznymi różnicami poziomu kompetencji kandydatów, 1% różnic w ocenie mógł być przypisany specyfice wykonywanego zadania, zaś 12% wynikało z indywidualnego sposobu oceniania kandydatów przez każdego z egzaminatorów (McManus i in., 2006): „(...) some examiners being more stringent and requiring a higher performance than other examiners who are more lenient. Although the problem has been known qualitatively for at least a century, we know of no previous statistical estimation of the size of the effect in a large-scale, high-stakes examination. Here we use FACETS to carry out a multi-facet Rasch modelling of the paired judgements made by examiners in the clinical examination (PACES)”. Ten niepożądany efekt różnic między egzaminatorami został przeanalizowany pod kątem wpływu na wynik ogólny egzaminu dla poszczególnych kandydatów. Okazało się, że naturalna predyspozycja do oceniania surowiej bądź łagodniej, miała wpływ na wynik egzaminu (mógł być on potencjalnie inny) zaledwie w przypadku 4,1% kandydatów. 2,6% kandydatów powinno było zdać egzamin, zaś 1,5% otrzymało wynik niesłusznie wyższy, który pozwolił im zaliczyć egzamin.

W przypadku egzaminu ustnego ważnymi czynnikami, które mogą mieć wpływ na wyniki, są płeć i pochodzenie etniczne kandydatów oraz egzaminatorów i wynikające z tego ryzyko niejawnej dyskryminacji. W 2007 roku przeprowadzono badania (Dewhurst i in., 2007), w których faktycznie wykazano, że istnieje niewielki efekt interakcji pomiędzy pochodzeniem etnicznym kandydatów i egzaminatorów. Zaobserwowano, że egzaminatorzy pochodzący z mniejszości łagodniej oceniali kandydatów również pochodzących z innych niż europejskie kręgi etniczne, ale nie zaobserwowano dyskryminacji ze strony egzaminatorów o pochodzeniu europejskim w stosunku do żadnej z grup kandydatów. Efekt ten występował tylko przy tych zadaniach, w których była oceniana komunikacja, natomiast nie wykazano istnienia tego efektu w zadaniach wymagających umiejętności klinicznych, co może wskazywać na pewne czynniki kulturowe, które ujawniają się tylko w sytuacji bezpośredniego kontaktu z osobami z podobnych kręgów kulturowych.

Z dalszych badań nad wpływem pochodzenia etnicznego i płci na wyniki egzaminu (Dewhurst i in., 2007) wynikało także, że z PACES najgorzej radzili sobie mężczyźni o pozaeuropejskim pochodzeniu etnicznym i to znacznie gorzej od grupy mężczyzn o europejskim pochodzeniu oraz kobiet z pozaeuropejskich kręgów etnicznych. Kobiety radziły sobie lepiej niż mężczyźni, przy czym efekt ten był przypisany nie czynnikom proceduralnym, a faktycznym umiejętnościom kandydatów. Podobne wyniki uzyskali także Bessant i inni (Bes

## 5. Ograniczenia/Uwagi

Egzamin MRCP(UK) musi być poddany skrupulatnej ocenie pod kątem wiarygodności nie tylko dlatego, że stanowi obowiązkowy etap na indywidualnej ścieżce kariery, lecz także dlatego, że lekarz jest zawodem zaufania publicznego i brak kompetencji może prowadzić do ryzyka po stronie pacjentów. Uwarunkowania polityczne i kulturowe sprawiły, że branża medyczna zapewniła sobie historycznie pewną niezależność w zakresie oceny kompetencji lekarzy świadczących usługi, prowadzenia szkoleń oraz udzielania licencji zawodowej. Prawo do takiego samostanowienia nabyła w zamian za zapewnienie wysokiej jakości świadczonych usług. Zaufanie społeczne, jakim została obdarzona branża medyczna w ostatnich kilkudziesięciu latach jednak zmalało, a jakość usług medycznych została poddana krytyce. Utrzymanie elementarnej niezależności college'ów wymaga zatem, by proces zarówno szkolenia lekarzy, jak i ich egzaminowania był transparentny i spełniał określone standardy.

Obu wspomnianym wyżej celom służyły zmiany wprowadzane do procedury przeprowadzania lub formuły MRCP(UK). Większość z nich była wprowadzana od lat sześćdziesiątych ubiegłego wieku w związku z daleko idącymi zmianami w myśleniu o egzaminach, które wynikały zarówno z postępu technologicznego, z dostępności mocy obliczeniowej komputerów, jak i z postępów w dziedzinie psychometrii. Przede wszystkim jednak zmiany te wynikały z postrzegania roli w społeczeństwie i obowiązków lekarza oraz rosnącego zrozumienia praw pacjenta.

Celem utrzymywania wysokich wymagań w stosunku do kandydatów MRCP(UK) jest zapewnienie jak najlepszej opieki lekarskiej pacjentom i utrzymanie wysokiego poziomu zaufania społecznego do branży medycznej. Wynikiem projektów badawczych dotyczących MRCP(UK) było między innymi potwierdzenie wysokiej zgodności ocen egzaminatorów odnośnie egzaminu PACES, brak dyskryminacji ze względu na pochodzenie, potwierdzenie, że wszystkie części MRCP(UK) przewidują wyniki innych egzaminów medycznych

i oceny w miejscu pracy oraz że słabe wyniki uzyskane na MRCP(UK) wskazują na późniejsze potencjalne problemy w wykonywaniu zawodu lekarza. W obecnej postaci egzamin wykorzystuje różne metody weryfikacji stosownych umiejętności, a jednocześnie jest on ustandaryzowany, rzetelny i trafny i spełnia międzynarodowe standardy ustalone dla testów edukacyjnych i psychologicznych. Duży nacisk położony w ZSK na rzetelność i trafność weryfikacji efektów uczenia się wymaganych dla kwalifikacji jest zatem zasadny i wskazany.

Na wyższym poziomie zgodność ze standardami oznacza, że MRCP(UK) realizuje swój cel, którym jest dopuszczenie do szkolenia specjalistycznego tych lekarzy, których wiedza, umiejętności i profesjonalne podejście dobrze rokują na późniejsze wykonywanie zawodu lekarza, zgodnie ze standardami profesji zawartymi w *Good medical practice*. Mimo że samo postrzeganie tego, czym charakteryzuje się dobry lekarz, ulegało na przełomie ostatnich kilkudziesięciu lat wielu zmianom – co było podyktowane zarówno postępującą profesjonalizacją usług, jak i niejednokrotnie wydarzeniami medialnymi z udziałem niekompetentnych lekarzy – to jednak standardy wytyczone nowym pokoleniom medyków stawały się obiektywnie coraz wyższe. Oprócz rozbudowania zasad etyki zawodowej, wytycznych dotyczących relacji z kolegami i kładzenia coraz większego nacisku na ciągły rozwój kompetencji, szczególnej wagi nabrały umiejętności komunikacji z pacjentami i innymi osobami zaangażowanymi w opiekę nad chorymi, co wynikało bezpośrednio z potrzeb i zażaleń zgłaszanych przez samych chorych. Zmiana myślenia w tym zakresie doprowadziła do uwzględnienia wyżej wymienionych umiejętności w programie nauczania medycyny, a w przypadku MRCP(UK) zaowocowała zmianami wprowadzonymi na przykład do egzaminu PACES.

Dbanie o jakość egzaminu wymaga od college'u jako instytucji egzaminującej nie tylko monitorowania całego procesu kształcenia lekarzy, ale także szeroko zakrojonej współpracy z innymi podmiotami, w tym tymi bezpośrednio odpowiedzialnymi za szkolenie oraz tymi, które sprawują nadzór nad całym procesem kształcenia. Te dwie funkcje, jak pokazano, są w Wielkiej Brytanii rozdzielone. Pozwala to na większą transparentność systemu kształcenia oraz procesów egzaminowania i zapewniania jakości tych usług, a przez to także podwyższa zaufanie do całej branży.