

24.11.2017

Autoreferat w języku polskim

Dr n. farm. Mariusz Panczyk

Załącznik nr 2

1. DANE OSOBOWE

Imię i nazwisko: Mariusz Panczyk

Aktualne miejsce pracy: Zakład Dydaktyki i Efektów Kształcenia
Wydział Nauki o Zdrowiu
Warszawski Uniwersytet Medyczny

2. POSIADANE DYPLOMY, STOPNIE NAUKOWE

2003 Dyplom ukończenia jednolitych studiów magisterskich na kierunku farmaceutycznym Uniwersytetu Medycznego w Łodzi w zakresie technologicznym (tytuł zawodowy magistra farmacji)

2007 Dyplom doktora nauk farmaceutycznych w zakresie biologii molekularnej uzyskany na Wydziale Farmaceutycznym Uniwersytetu Medycznego w Łodzi, na podstawie rozprawy:

Polimorfizm genu MDR1 (ABCB1) kodującego białko oporności wielolekowej u osób chorych na raka jelita grubego w populacji łódzkiej

Promotor:

prof. dr hab. Marek Mirowski

Recenzenci:

prof. dr hab. Wanda Baer-Dubowska

prof. dr hab. Jacek Bartkowiak

2014 Dyplom ukończenia studiów podyplomowych na Wydziale Prawa i Administracji Uniwersytetu Warszawskiego w zakresie ochrony własności intelektualnej

3. UKOŃCZONE KURSY I SZKOLENIA

- 2017
- *EndNote online I – tworzenie i zarządzanie biblioteką rekordów* (Clarivate Analytics; Warszawa, 17 stycznia 2017 r.);
 - *EndNote online II – wykorzystanie w pisaniu publikacji naukowych* (Clarivate Analytics; Warszawa, 24 stycznia 2017 r.);
 - *Jak zadbać o statystyczną istotność wyników badań empirycznych – moc testu i liczebność próby?* (StatSoft Polska; Kraków, 7 lutego 2017 r.);
 - *Jak w praktyce korzystać z zaawansowanej analizy danych?* (StatSoft Polska; Kraków, 16 lutego 2017 r.);
 - *Jak wybrać czasopismo do publikacji mojej pracy?* (Clarivate Analytics; Warszawa, 17 lutego 2017 r.);
 - *Jak poprawnie stosować, prezentować i interpretować wyniki analizy przeżyć?* (StatSoft Polska; Kraków, 9 marca 2017 r.);
 - *Jak autorzy publikacji naukowych prezentują i interpretują statystyczne wyniki badań?* (StatSoft Polska; Kraków, 21 marca 2017 r.);
 - *Wykrywanie fraud detection z wykorzystaniem analizy danych* (StatSoft Polska; Kraków, 6 czerwca 2017 r.);
 - *Text mining i web scrapping w R* (Uniwersytet Warszawski; Warszawa, 13 maja–24 czerwca 2017 r.);
 - *Genomika i bioinformatyka* (MBS Serwis dla Biologii Molekularnej; Warszawa, 15–17 września 2017 r.);
 - *Web of Science Big Data* (Clarivate Analytics; Warszawa, 18 października 2017 r.);
 - *Web of Science – baza patentów Derwent Innovations Index* (Thomson Reuters; Warszawa, 8 września 2016 r.);
 - *Statystyka matematyczna od podstaw w programie R* (Uniwersytet Warszawski; Warszawa, 19 grudnia 2015–29 stycznia 2016 r.);
 - *OSCE – Objective Structured Clinical Examination, nowoczesne formy wersyfikacji efektów kształcenia praktycznego w oparciu o zaawansowane systemy symulatorów medycznych* (ProLongaVita; Międzyzdroje, 26–29 marca 2015 r.);
 - *How to use the Cochrane Library in your daily practice* (John Wiley & Sons, Inc.; Warszawa, 17 września 2014 r.);
 - *Essential Science Indicators na platformie InCites* (Thomson Reuters; Warszawa, 29 września 2015 r.);
 - *Nowe Journal Citation Reports na platformie InCites* (Thomson Reuters; Warszawa, 30 września 2015 r.);
 - *Web of Science Next Generation* (Thomson Reuters; Warszawa, 29 kwietnia 2014 r.);
 - *Impact Factor i czasopisma w bazie Journal Citation Reports* (Thomson Reuters; Warszawa, 30 października 2014 r.);
 - *Selekcja czasopism do bazy Web of Science Core Collection* (Thomson Reuters; Warszawa, 30 października 2014 r.);

- *Szkolenie w zakresie doskonalenia dydaktycznego*
(Biuro ds. Jakości Kształcenia Warszawski Uniwersytet Medyczny; Warszawa, 19 listopada 2014 r.);
- 2012 ○ *Przygotowanie zgłoszeń oraz procedury ochrony wynalazków i wzorów użytkowych*
(Wyższa Szkoła Zawodowa Kosmetyki i Pielęgnacji Zdrowia; Warszawa, 29 listopada 2012 r.);
- *Modernizacja kształcenia zawodowego na podstawie modułowych programów nauczania*
(Mazowieckie Samorządowe Centrum Doskonalenia Nauczycieli, Warszawa 26 maja–2 czerwca 2010 r.);
- *Gene Expression School*
(Applied Biosystems; Warszawa, 16 czerwca 2010 r.);
- *Emisja i higiena głosu, czyli o tym, jak mówić długo i donośnie, zupełnie się nie męcząc*
(ABC poprawnej wymowy; Warszawa, 29 października 2010 r.);
- 2009 ○ *Ochrona danych osobowych*
(BINF Bezpieczeństwo Informacji; Warszawa, 26 listopada 2009 r.);
- *Badania kliniczne oraz medyczne badania obserwacyjne, organizacja, realizacja*
(Centrum Medycznego Kształcenia Podyplomowego; Warszawa, 16–17 lutego 2008 r.);
- *Kurs pedagogiczny*
(Akademia Pedagogiki Specjalnej i Warszawski Uniwersytet Medyczny; Warszawa, marzec–maj 2008 r.);
- *Statystyczna metodologia badań klinicznych*
(Centrum Medycznego Kształcenia Podyplomowego; Warszawa, 9–13 czerwca 2008 r.);
- *Metody i instrumenty nowoczesnej dydaktyki medycznej*
(Centrum Medycznego Kształcenia Podyplomowego; Warszawa, 18–19 września 2008 r.).

4. INFORMACJE O DOTYCHCZASOWYM ZATRUDNIENIU W JEDNOSTKACH NAUKOWYCH

Urodziłem się 22 stycznia 1979 roku w Łodzi. W 1998 roku ukończyłem XXXI Liceum Ogólnokształcące im. Ludwika Zamenhofa w Łodzi, a w 2003 roku studia magisterskie na kierunku farmaceutycznym Uniwersytetu Medycznego w Łodzi. W trakcie studiów podjąłem działalność naukową w ramach Studenckiego Koła Naukowego działającego przy Pracowni Biologii Molekularnej i Farmakogenomiki Zakładu Biochemii Farmaceutycznej UM w Łodzi. Efektem tej działalności była praca magisterska pt. „Polimorfizm genu *MDR1* oraz ocena ekspresji genu *BCRP* w ostrych i przewlekłych białaczkach osób dorosłych”.

W latach 2003–2007 w ramach Studium Doktoranckiego realizowałem pracę doktorską pod kierunkiem prof. dra hab. Marka Mirowskiego. Na realizację projektu badawczego uzyskałem finansowanie z funduszy Komitetu Badań Naukowych (KBN/2P05F 02628) oraz Marszałka Województwa Łódzkiego. W roku 2007 uzyskałem z wyróżnieniem dyplom doktora nauk farmaceutycznych w zakresie biologii molekularnej, na podstawie rozprawy „Polimorfizm genu *MDR1* (*ABCB1*) kodującego białko oporności wielolekowej u osób chorych na raka jelita grubego w populacji łódzkiej”, której recenzentami byli prof. dr hab. Wanda Baer-Dubowska oraz prof. dr hab. Jacek Bartkowiak. Od września do grudnia 2007 roku odbywałem staż zawodowy jako specjalista ds. rejestracji leków w SciencePharma Sp. z o.o. w Warszawie.

Od 2008 roku do dnia dzisiejszego jestem związany zawodowo z Wydziałem Nauki o Zdrowiu Warszawskiego Uniwersytetu Medycznego (dawniej Akademia Medyczna w Warszawie). Pracuję jako adiunkt (od 2015 roku jako asystent) w Zakładzie Dydaktyki i Efektów Kształcenia, w którym od maja 2014 roku kieruję Pracownią Ewaluacji i Doskonalenia Kształcenia Medycznego.

5. WSKAZANIE OSIĄGNIĘCIA NAUKOWEGO wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003 roku o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. z 2003 r. Nr 65, poz. 595 z późn. zm.)

Podstawą do ubiegania się o stopień doktora habilitowanego jest dzieło opublikowane w całości (monografia).

5.1. Tytuł osiągnięcia naukowego:

„Ocena systemu kwalifikacji kandydatów oraz efektywności nauczania na kierunku pielęgniarstwo studiów I i II stopnia na Warszawskim Uniwersytecie Medycznym w latach 2005–2015”

5.2. Autor/tytuły publikacji, nazwa wydawnictwa, miasto, rok wydania:

Panczyk M. „Ocena systemu kwalifikacji kandydatów oraz efektywności nauczania na kierunku pielęgniarstwo studiów I i II stopnia na Warszawskim Uniwersytecie Medycznym w latach 2005–2015”. Wydawnictwo Warszawskiego Uniwersytetu Medycznego: Warszawa, 2017.

5.3. Recenzje wydawnicze przygotowali:

- dr hab. Izabella Uchmanowicz, prof. nadzw.
Zakład Pielęgniarstwa Internistycznego
Katedra Pielęgniarstwa Klinicznego
Wydział Nauk o Zdrowiu, Uniwersytet Medyczny we Wrocławiu
- dr hab. Elżbieta Grochans, prof. nadzw.
Zakład Pielęgniarstwa
Wydział Nauk o Zdrowiu, Pomorski Uniwersytet Medyczny

5.4. Omówienie celu naukowego ww. pracy i osiągniętych wyników wraz z omówieniem ich ewentualnego wykorzystania:

5.4.1. Wprowadzenie i cel naukowy

Stały rozwój medycyny, wzrost długości życia i wynikające z tego rosnące koszty w ochronie zdrowia sprawiły, że w większości krajów, w tym także w Polsce, konieczne stało się zwiększenie efektywności istniejących systemów zdrowotnych. Jednym z założeń reformy systemu ochrony zdrowia było dostosowanie modelu kształcenia kadr medycznych do potrzeb społecznych i wyzwań demograficznych. Aby możliwa była realizacja nowej koncepcji funkcjonowania ochrony zdrowia, w systemie kształcenia pielęgniarek wprowadzono niezbędne zmiany organizacyjne i programowe. Działania te były związane z koniecznością przygotowania wykwalifikowanej kadry gwarantującej wysoką jakość świadczonych usług medycznych. Społeczeństwo oczekuje rzetelnej wiedzy i kwalifikacji od osób świadczących opiekę zdrowotną, a jedynie kompetentna i wykształcona kadra może takim wymogom sprostać. Transformacja szkolnictwa pielęgniarstwa była częścią kompleksowej strategii działań, zmierzającej do zapewnienia odpowiedniej liczby profesjonalnie wykształconych pielęgniarek.

Pomimo wzrastającej liczby absolwentów z dyplomem licencjata pielęgniarstwa oraz magistrów pielęgniarstwa prognozy na najbliższe lata nie są optymistyczne [1]. W latach 2007–2009 mniej niż połowa absolwentów zgłaszała się do odpowiednich Okręgowych Izb Pielęgniarek i Położnych w celu otrzymania prawa wykonywania zawodu. Szacunki Naczelnej Izby Pielęgniarek i Położnych mówią, że do roku 2020 w Polsce zabraknie 60 860 pielęgniarek. W stosunku do roku 2009 oznacza to ubytek kadr o 33% [2]. Powyższe dane skłaniają do poszukiwania rozwiązań zapewniających możliwość kształcenia dużej liczby studentów pielęgniarstwa zmotywowanych do podjęcia pracy w zawodzie po skończeniu studiów. Zadanie to wymaga podjęcia wysiłków w kierunku poprawy istniejących mechanizmów rekrutacji oraz przeglądu obecnie funkcjonujących zasad naboru kandydatów na studia na ten kierunek. Dobór kandydatów o najlepszych cechach i predyspozycjach może zwiększyć szanse, że po skończeniu studiów podejmą oni pracę zgodnie z wyuczonym zawodem. Wzrost motywacji studentów także można osiągnąć – przez przygotowanie adekwatnych do oczekiwań rynku pracy atrakcyjnych programów studiów, dbanie o dobrą jakość dydaktyki oraz przyjazne studentom warunki studiowania. Niektóre z tych zadań mogą być realizowane przez uczelnię we własnym zakresie, podczas gdy inne wymagają ścisłej współpracy między uczelniami a samorządem zawodowym i towarzystwami naukowymi, tak aby wspólnie wypracować działania podnoszące jakość i efektywność kształcenia nowych pielęgniarek.

Założenie:

Dbłość o wysokie standardy kształcenia i efektywne nauczanie w tak ważnym dla systemu ochrony zdrowia zawodzie, jak pielęgniarstwo, powinno iść w parze ze stałą kontrolą mechanizmów kwalifikacji kandydatów na ten kierunek studiów. Nawet najlepsze systemy zapewniające jakość kształcenia nie spełnią swojej funkcji, jeśli pewna grupa studiujących nie będzie w stanie sprostać wymaganiom ze względu na brak odpowiednich cech i predyspozycji, które nie zostały prawidłowo ocenione i zweryfikowane podczas naboru [3-5]. Strategia Evidence-Based Admission Policy zakłada, że polityka rekrutacyjna uczelni opiera się na sprawdzaniu efektywności systemu kwalifikacji kandydatów na studia przez monitorowanie postępów studenta poczynionych w toku studiów, a także na śledzeniu losów absolwentów uczelni. Dowody zbierane z takich obserwacji pozwalają na doskonalenie trafności i rzetelności kryteriów selekcji kandydatów na studia.

Cel główny:

Określenie czynników, które miały istotny wpływ na jakość procesu rekrutacji kandydatów na studia oraz wyniki kształcenia studentów na studiach stacjonarnych pierwszego i drugiego stopnia prowadzonych na kierunku pielęgniarstwo Wydziału Nauki o Zdrowiu Warszawskiego Uniwersytetu Medycznego (WUM) w latach 2005–2015.

Cele szczegółowe:

1. Ocena przebiegu procesu selekcji kandydatów na stacjonarne studia pierwszego i drugiego stopnia na kierunku pielęgniarstwo z uwzględnieniem zmian, jakie zachodziły w naborze w latach 2005–2015.
2. Oszacowanie trafności prognostycznej kryteriów naboru na stacjonarne studia pierwszego i drugiego stopnia na kierunku pielęgniarstwo.
3. Analiza spójności systemu oceniania osiągnięć edukacyjnych studentów oraz jego zgodności z założeniami zawartymi w kierunkowych standardach kształcenia.
4. Wytypowanie istotnych determinantów sukcesu/porażki akademickiej studentów, którzy podjęli kształcenie na studiach pierwszego stopnia na kierunku pielęgniarstwo.
5. Poszukiwanie modelu prognostycznego pozwalającego przewidywać wyniki kształcenia studentów studiów drugiego stopnia w oparciu o metodologię szacowania edukacyjnej wartości dodanej.

5.4.2. Metodologia przeprowadzonych badań

OCENA PRZEBIEGU PROCESU SELEKCJI KANDYDATÓW NA STUDIA

Dobrze zaplanowana strategia i polityka rekrutacyjna pozwalają na doskonalenie wcześniej przyjętych zasad i dostosowywanie mechanizmów selekcji do zmieniających się warunków, takich jak: liczba kandydatów oraz ich ogólny poziom wiedzy i umiejętności na wejściu, a także potrzeby kadrowe w zakresie danego zawodu. Jeśli chcemy prowadzić spójną, opartą na dowodach politykę rekrutacyjną (Evidence-Based Admission Policy), to powinniśmy mieć możliwość walidacji narzędzi stosowanych w trakcie kwalifikacji kandydatów.

Według proponowanej metodologii proces walidacji powinien przebiegać w cyklu, na który składają się następujące etapy: (I) planowanie, (II) zastosowanie, (III) obserwacja, (IV) krytyczna ocena i (V) wnioskowanie ze wskazaniem nowych wytycznych. Zamknięcie cyklu prowadzi do rozpoczęcia całego procesu od początku (rewalidacja) w celu określenia, czy po wdrożeniu nowych rekomendacji osiągnięto spodziewane korzyści [6]. Dwa z pierwszych wymienionych powyżej etapów walidacji są każdego roku przeprowadzane wraz z kolejnym naborem kandydatów. Faza planowania jest częścią przygotowywania uchwały rekrutacyjnej, która następnie każdorazowo podlega zatwierdzeniu przez senat uczelni. Etap zastosowania wiąże się natomiast z realizacją procedury naboru, na który składa się między innymi: gromadzenie danych dotyczących kandydatów, obliczenie punktów kwalifikacyjnych i utworzenie listy rankingowej osób aplikujących na studia. Kolejne trzy etapy walidacji – obserwacja, krytyczna ocena i wnioskowanie, są przedmiotem opisywanej analizy i zostaną zaprezentowane w kolejnej części autoreferatu.

TRAFNOŚĆ PROGNOSTYCZNA KRYTERIÓW NABORU

Zgodnie z koncepcją Evidence-Based Admission Policy decyzja dotycząca przyjęcia/nieprzyjęcia na studia powinna mieć oparcie w ustaleniach poczynionych na podstawie analizy predykcyjnej danych empirycznych. Według Grega Perfetta przy określaniu potencjalnych obiektywnych czynników predykcyjnych warunkujących osiągnięcie przyszłego sukcesu akademickiego/zawodowego przez kandydata na studia należy się kierować czterema podstawowymi zasadami: (I) rzetelność, (II) trafność, (III) stabilność w czasie, (III) stosowalność w praktyce akademickiej [7]. Wobec powyższego, dopasowanie kryteriów kwalifikacyjnych wiąże się między innymi z oceną trafności prognozy, którą należy rozumieć jako stopień zgodności przewidywań przyszłych wyników studenta na podstawie pomiaru wstępnych cech kandydata na studia [8].

SPÓJNOŚĆ SYSTEMU OCENIANIA OSIĄGNIĘĆ EDUKACYJNYCH STUDENTÓW

Obok trafnej i rzetelnej selekcji kandydatów, istotnym czynnikiem wpływającym na efektywność kształcenia nowej kadry pielęgniarskiej jest jakość przygotowania zawodowego. Jakość ta zależy w dużym stopniu od realizacji programu studiów i jego zgodności z obowiązującymi standardami kształcenia oraz wytycznymi Krajowej Rady Akredytacyjnej Szkół Pielęgniarek i Położnych. Rada w swojej uchwale wśród rekomendacji dotyczących sposobu realizacji programu kształcenia wymienia, że *program studiów dla kierunku pielęgniarstwo określa sposoby weryfikacji zakładanych efektów kształcenia osiągniętych przez studenta* [9]. Ta sama uchwała wskazuje ponadto, że w zakresie oceny jakości kształcenia i osiągnięć uczelni *wewnętrzny system zapewnienia jakości kształcenia umożliwia prowadzenie działań na rzecz doskonalenia programu kształcenia na kierunku studiów pielęgniarstwo, a w szczególności przewiduje odpowiednie formy weryfikowania wszystkich efektów kształcenia w zakresie wiedzy, umiejętności oraz kompetencji społecznych* [9].

Problematyka oceniania w edukacji medycznej oraz kształceniu specjalistów nauk o zdrowiu jest przedmiotem wielu badań i budzi szerokie zainteresowanie środowiska akademickiego [10-12]. Fakt ten nie jest zaskakujący, ponieważ ocena studentów jest postrzegana jako jeden z ważniejszych elementów całego systemu edukacyjnego. Z jednej strony określa ona stopień osiągnięcia przez studenta zakładanych efektów kształcenia, a z drugiej może być także miarą jakości procesu edukacyjnego [11]. Bez względu na cel, jakim służy ocena, zawsze wiąże się ona z mniej lub bardziej systematycznym gromadzeniem danych z obserwacji, które prowadzą do postawienia wniosków na temat cech ocenianego studenta. Aby proces ten był wysoce obiektywnym źródłem informacji na temat osiągniętych efektów kształcenia, powinien on spełniać pewne kryteria, określane jako cechy diagnozy edukacyjnej. Do właściwości pomiaru edukacyjnego, które pozwalają na jego ocenę i optymalizację, należą przede wszystkim: obiektywizm, rzetelność i trafność [12-15].

DETERMINANTY SUKCESU/PORAŻKI AKADEMICKIEJ

Kolejnym ważnym problemem obok konieczności zapewnienia dobrej jakości selekcji kandydatów oraz odpowiedniej efektywności funkcjonowania systemu oceniania postępów edukacyjnych studentów jest wysoki odsetek niepowodzeń wśród studentów, którzy podjęli studia. Problem ten jest dostrzegany w wielu krajach [16-20]. W Polsce dotychczas nie przeprowadzono systematycznych badań obserwacyjnych, które pozwoliłyby na oszacowanie rozmiarów tego zjawiska oraz zidentyfikowanie głównych przyczyn niepowodzeń/rezygnacji w trakcie

studiowania. Brak tego typu badań uniemożliwia podjęcie skutecznych działań wspierających proces naboru odpowiednich kandydatów na studia, co zmniejszałoby ryzyko niepowodzeń z powodu złego dopasowania kryteriów kwalifikacyjnych [21].

Teoretyczne modele sukcesu/porażki, budowane na podstawie różnych koncepcji socjologicznych i psychologicznych, pozwalają na projektowanie badań mających dostarczyć empirycznych dowodów, które wspierałyby określone kierunki działań, jakie można podjąć, aby zwiększyć u studiujących odsetek sukcesów. Wnioski i rekomendacje z tego rodzaju badań pozwalają wdrażać odpowiednie rozwiązania administracyjne związane z rekrutacją na studia i obsługą studentów, tworzyć warunki ułatwiające proces socjalizacji i adaptacji w środowisku akademickim oraz wspierać wybrane osoby, dla których ryzyko niepowodzenia jest szczególnie wysokie [22, 23].

W światowym piśmiennictwie odnajdujemy co najmniej kilka teorii i modeli, które twórcy próbują usystematyzować i wyjaśnić złożony problem sukcesu/porażki w edukacji akademickiej. Wielu badaczy zajmujących się tą problematyką w kontekście kształcenia pielęgniarek jako ramy teoretyczne w swoich pracach wykorzystuje najczęściej koncepcję opisaną przez Vincenta Tintę i/lub Johna Beana [23-31]. Jednakże w pierwszej dekadzie XXI wieku pojawiły się nowe modele opisujące teoretyczne podstawy sukcesu/porażki, opracowane *stricte* na potrzeby badań dotyczących edukacji studentów pielęgniarstwa [26, 32, 33].

W 2003 roku Elisabeth Shelton opublikowała pracę na podstawie wyników eksperymentalnych badań dotyczących możliwości przeciwdziałania niepowodzeniom edukacyjnym wśród studentów pielęgniarstwa [26]. Bardziej zaawansowaną koncepcję, także dotyczącą wyników kształcenia pielęgniarek, zaproponowała Marianne Jeffreys [32, 33]. Model NURS (*Nursing Universal Retention and Success*) został stworzony w oparciu o założenia Beana [34, 35] i Bandury [36] oraz liczne dane empiryczne, które umożliwiły stworzenie listy skategoryzowanych czynników mogących mieć istotny wpływ na sukces studenta w toku kształcenia na kierunku pielęgniarstwo, zarówno na studiach pierwszego i drugiego stopnia, jak też na studiach doktoranckich (studia trzeciego stopnia) czy w ramach kształcenia specjalizacyjnego [32, 33]. Do modelu NURS wprowadzono wiele zmiennych, dla których istnieją wiarygodne przesłanki co do ich wpływu na sukces/porażkę podczas edukacji.

Tradycyjnie w ocenie sukcesu/porażki jako predyktory testuje się: (1) wyniki kwalifikacji na studia, (2) terminowe ukończenie pierwszego roku studiów lub całego cyklu kształcenia, (3) średnią ocen wyliczaną za kolejne lata studiów lub za poszczególne bloki programowe, (4) ocenę uzyskaną na egzaminie dyplomowym

lub zawodowym, (5) zdobycie zatrudnienia w zawodzie po ukończeniu studiów, (6) uzyskanie specjalizacji oraz (7) zajmowane stanowiska [37].

W Polsce dostępna jest bardzo ograniczona liczba danych empirycznych o podłożach sukcesu/porażki w edukacji. W polskich warunkach kształcenia pielęgniarek nie dokonano dotychczas oceny trafności modeli teoretycznych stworzonych w Stanach Zjednoczonych. Można jednakże przyjąć założenie, że niektóre spośród opisanych modeli są na tyle uniwersalne, że z powodzeniem mogą być zastosowane w badaniach krajowych. Dlatego też w analizie determinantów sukcesu/porażki studentów, którzy podjęli kształcenie na studiach pierwszego stopnia na kierunku pielęgniarstwo na WUM, wykorzystano założenia teoretyczne modelu NURS.

SZACOWANIE EDUKACYJNEJ WARTOŚCI DODANEJ

Ostatnim elementem opisywanych badań było poszukiwanie najlepszego modelu prognostycznego, który pozwalałby z dużą dozą pewności przewidywać przyszłe wyniki kształcenia studentów. W tym celu posłużono się metodologią szacowania edukacyjnej wartości dodanej (EWD). Koncepcja EWD pojawiła się w połowie lat 70., jednak większe zainteresowanie praktycznym zastosowaniem takiej analizy w pomiarze dydaktycznym służącym obiektywnej ocenie postępów uczniów nastąpiło dopiero w ciągu ostatnich 25 lat [38]. EWD jest definiowana jako przyrost wiedzy studentów zachodzący w wyniku prowadzenia procesu edukacyjnego [39]. EWD rozumiana wprost odnosi się do idei postępów studentów w nauce. W idei tej w ocenie efektywności kształcenia uwzględnia się stan początkowy, zwany potencjałem edukacyjnym. W szerszym znaczeniu EWD może być szacowana z uwzględnieniem czynników dodatkowych, dotyczących w szczególności zasobów (kadrowych, finansowych etc.), którymi dysponuje szkoła bądź instytucja akademicka [38].

W badanym okresie na kierunku pielęgniarstwo, zarówno na studiach pierwszego, jak i drugiego stopnia, nie stosowano standaryzowanych narzędzi pomiaru dydaktycznego. W związku z tym do oceny EWD na WUM można wykorzystać jedynie wskaźniki względne. Przyjmujemy, że wyniki uzyskane w ramach przedmiotów podstawowych (szczebel podstawowy) są odzwierciedleniem potencjału edukacyjnego studentów. Następnie używamy tej miary jako predyktora dla przewidywania osiągnięć studentów uzyskanych w toku kształcenia kierunkowego, klinicznego, specjalistycznego (szczebel wyższy). Oczekiwany wynik kształcenia dla wyższego szczebla odnosimy do faktycznie uzyskanych wyników, a różnica między tymi dwiema wartościami stanowi oszacowaną EWD. Wyliczona z tych wskaźników średnia dla danej uczelni może być miarą efektywności kształcenia [38, 40].

ZASTOSOWANE METODY ANALIZY DANYCH

1. OCENA PRZEBIEGU PROCESU SELEKCJI KANDYDATÓW NA STUDIA

W ocenie zdolności selekcyjnej stosowanych zasad naboru na studia posłużono się trzema metodami statystycznymi:

- analizą interakcji między zmiennymi niemetrycznymi w modelu logarytmiczno-liniowym [41],
- analizą dyskryminacyjną kryteriów rekrutacyjnych z wykorzystaniem krzywych ROC (*Receiver Operating Characteristics*) dla dwóch kategorii: przyjęty/nieprzyjęty oraz modyfikacją statystyki U zaproponowaną przez Hanleya i Hajian-Tilakiego do porównania AUC (*Area Under the Curve*) [42, 43],
- oceną predykcyjną wybranych zmiennych w modelu regresji logistycznej dla dychotomicznej zmiennej wynikowej: przyjęty/nieprzyjęty [44].

2. TRAFNOŚĆ PROGNOSTYCZNA KRYTERIÓW NABORU

Trafność prognostyczna to ocena zgodności wyników jednego pomiaru (w tym przypadku: wyników kwalifikacji na studia) z kryterium zewnętrznym, czyli osiągnięciami edukacyjnymi studentów mierzonymi średnią ocen. Kryterium zewnętrzne odnosi się do pomiaru edukacyjnego całkowicie niezależnego od oceny kompetencji dokonywanej podczas rekrutacji na studia. Trafność prognostyczna jest odmianą trafności kryterialnej, w której kryterium zewnętrzne jest oddalone w czasie od ocenianego pomiaru [45]. Do oszacowania trafności zastosowano metody korelacyjne i bardziej złożone modele regresji wielorakiej oraz analizę kowariancji (*analysis of covariance*, ANCOVA).

3. SPÓJNOŚĆ SYSTEMU OCENIANIA OSIĄGNIĘĆ EDUKACYJNYCH STUDENTÓW

Do oszacowania właściwości psychometrycznych systemu oceny osiągnięć studentów pielęgniarstwa wykorzystano metody analizy rzetelności, trafności i jednowymiarowości poszczególnych grup przedmiotów [13-15].

W ocenie rzetelności wykorzystano podejście polegające na analizie spójności wewnętrznej pomiaru osiągnięć studentów, stosując formułę zaproponowaną przez Lee Cronbacha [46]. Zgodnie z kryterium Juma Nunnally'ego jako dopuszczalny próg rzetelności dla alfa-Cronbacha przyjęto wartość przekraczającą 0,70 [46]. Wyznaczono ponadto dla każdego przedmiotu wartość indeksu dyskryminacyjnego w celu oszacowania, w jakim stopniu oceny z danego przedmiotu wpływają na całkowitą zgodność pomiaru osiągnięć studentów. Wartość progową dla tego indeksu przyjęto na poziomie 0,40 [47].

Oszacowanie trafności systemu oceny osiągnięć studentów przeprowadzono, wykorzystując dwa różne podejścia analityczne:

- Szacowanie trafności teoretycznej z wykorzystaniem eksploracyjnej analizy czynnikowej [48];
- Szacowanie trafności kryterialnej, która opierała się na określeniu stopnia korelacji między wybranymi przedmiotami oraz grupami przedmiotów. Do weryfikacji powyższego założenia wyliczono wartości współczynników korelacji r-Spearmana.

Analizę jednowymiarowości każdej grupy przedmiotów przeprowadzono za pomocą analizy głównych składowych. Przyjęto, że za jednowymiarową uznaje się grupę przedmiotów, która spełnia kryterium Henry'ego Kaisera [49].

4. DETERMINANTY SUKCESU/PORAŻKI AKADEMICKIEJ

Do oceny czynników predykcyjnych, które mają potencjalny wpływ na ryzyko nieukończenia studiów (porażka akademicka), wykorzystano metodę analizy regresji logistycznej z nieliniową estymacją funkcji.

5. SZACOWANIE EDUKACYJNEJ WARTOŚCI DODANEJ

Do szacowania EWD zastosowano w pierwszej kolejności analizę korelacji r-Pearsona i liniową regresję jednozmienną z estymacją metodą najmniejszych kwadratów. Natomiast dla wieloletniego szacowania EWD ze zmiennymi kontrolnymi użyto metody regresji wielorakiej z estymacją grzbietową i postępującym krokowym wprowadzaniem zmiennych do testowanego modelu. Do ocenianych predyktorów należały zarówno zmienne socjodemograficzne, jak i kryteria kwalifikacji stosowane podczas naboru na studia pierwszego i drugiego stopnia. Do modelu włączono także oceny studentów z 16 egzaminów przedmiotowych przewidzianych programem studiów pierwszego stopnia oraz średnie ocen wyliczone po pierwszym, drugim i trzecim roku studiów przeddyplomowych.

5.4.3. Uzyskane wyniki

STUDIA PIERWSZEGO STOPNIA

1. OCENA PRZEBIEGU PROCESU SELEKCJI KANDYDATÓW NA STUDIA STACJONARNE Z UWZGLĘDNIENIEM ZMIAN, JAKIE ZACHODZIŁY W NABORZE W LATACH 2005–2015

Zgodnie ze stosowanymi przez WUM zasadami rekrutacji podstawą przyjęć na kierunek pielęgniarstwo studia pierwszego stopnia są wyniki maturalne uzyskane przez kandydata z trzech przedmiotów (język polski, język obcy i przedmiot dodatkowy). Taki system selekcji nie jest więc oparty na specyficznym dopasowaniu kryteriach oceniających jakieś szczególne cechy czy predyspozycje kandydata. Stosowane w Polsce rozwiązania są wyraźnie odmienne od światowych trendów w zakresie dostosowywania polityki przyjęć do specyficznych potrzeb związanych z kształceniem studentów pielęgniarstwa [50-52]. Można założyć, że funkcjonujący w kraju system naboru na studia jest pewnym rozwiązaniem pośrednim między strategią wyboru kandydatów o ściśle określonych predyspozycjach (jak np. w systemie amerykańskim) a polityką rekrutacyjną, w której dopiero po pierwszym roku studiów dokonujemy właściwej selekcji studentów na dany kierunek (jak np. w systemie francuskim) [53].

Warto wspomnieć, że przed ogólnopolską reformą oświaty i szkolnictwa wyższego, która wprowadziła wymóg uwzględniania podczas rekrutacji na studia przedmiotowych wyników z egzaminu maturalnego, uczelnie medyczne wykorzystywały do selekcji kandydatów na kierunki lekarski, lekarsko-dentystyczny, farmację i analitykę medyczną specjalnie przygotowane egzaminy testowe. Stosowany do 2005 roku ogólnopolski egzamin wstępny był opracowywany przez ekspercki zespół specjalistów wywodzących się bezpośrednio ze środowiska akademii medycznych. Na potrzeby kwalifikacji kandydatów corocznie był przygotowywany specyficznym dopasowany do danego kierunku studiów egzamin testowy, który odrębnie oceniał predyspozycje kandydata, w zależności od tego, czy rekrutacja dotyczyła studiów lekarskich czy farmaceutycznych. Wprowadzenie nowych i obecnie obowiązujących rozwiązań prawnych skutecznie ograniczyło możliwości prowadzenia tego typu naboru na studia wyższe w Polsce.

Na podstawie uzyskanych wyników analizy dyskryminacyjnej oraz przy uwzględnieniu powyższych zastrzeżeń można przyjąć, że generalnie dotychczasowe zasady naboru kandydatów na kierunek pielęgniarstwo studia pierwszego stopnia dobrze spełniły swoje zadanie. Przyjęci kandydaci uzyskiwali każdego roku średnio wyższe wyniki z matury (język polski i angielski oraz biologia) niż ogół maturzystów

w Polsce. Świadczy to o tym, że osoby podejmujące naukę na WUM są generalnie lepiej przygotowane do studiów wyższych niż przeciętny maturzysta w Polsce. Dobrej kwalifikacji sprzyjał ponadto duży zakres zmienności punktacji rankingowej, który w połączeniu ze znaczną liczbą aplikujących na studia pozwalał na dobre różnicowanie kandydatów, co z kolei zwiększało rzetelność całego procesu. Poza tym wprowadzenie do zasad rekrutacji opcji wyboru przedmiotów premiowanych prawdopodobnie zwiększyło trafność selekcji, co pozwala zakładać, że wśród rozpoczynających studia nie było znacznego odsetka osób niedostatecznie przygotowanych do podjęcia nauki na studiach wyższych [54].

Z wyników analizy zmiennych socjodemograficznych z wykorzystaniem modelu regresji logistycznej można wnioskować, że w porównaniu z kandydatami, którzy zdawali maturę poza Warszawą, lepsze lokaty na listach rankingowych osiągnęli absolwenci warszawskich szkół średnich. Ten sam model regresji wskazuje także na lepsze prognozy dla kandydatów mających tzw. starą maturę. Obserwacja ta została także potwierdzona w modelowaniu logarytmiczno-liniowym, w ocenie interakcji zmiennych niemetrycznych. Nie wykazano natomiast istotnego wpływu płci oraz wieku kandydata na powodzenie podczas rekrutacji.

Szczegółowa ewaluacja zastosowanych metod kwalifikacji kandydatów, uzyskana w analizie dyskryminacyjnej, pozwoliła na prześledzenie pewnych niekorzystnych zmian związanych z jakością poszczególnych rekrutacji w okresie od roku akademickiego 2005/06 do 2015/16. Generalnie zaobserwowano zmniejszenie zdolności selekcyjnej trzech kryteriów rekrutacyjnych, co jest szczególnie zauważalne w ostatnich sześciu latach. Ma to zapewne związek z niższym demograficznym, a co za tym idzie – zmniejszeniem ogólnej liczby kandydatów na studia. W sytuacji mniejszej konkurencji absolwenci szkół średnich z dobrymi i bardzo dobrymi wynikami egzaminów maturalnych mają do wyboru na rynku edukacyjnym „bardziej atrakcyjne” kierunki studiów niż pielęgniarstwo. Taki stan rzeczy przyczynia się w konsekwencji do zmniejszenia dostępnej puli osób o dobrym przygotowaniu do studiowania, utrudnia różnicowanie kandydatów i zwiększa ryzyko niewłaściwego doboru studentów oraz prawdopodobieństwo niepowodzenia w toku studiów.

Stosowane na WUM warunki przyjęć na kierunek pielęgniarstwo studia stacjonarne pierwszego stopnia, nie licząc punktów dodatkowych za nadprogramowe kwalifikacje (np. posiadanie dyplomu zawodowego zgodnego z wybranym kierunkiem studiów), są w całości oparte na ocenie tzw. cech poznawczych (*cognitive*). Jednak wielu współczesnych badaczy zajmujących się doбором kryteriów kwalifikacji kandydatów na studia w uczelniach medycznych zwraca uwagę na konieczność holistycznej oceny kompetencji wejściowych. Wielowymiarowe podejście do oceny kandydatów, postulowane przez środowisko

akademickie, zakłada stosowanie kryteriów poznawczych oraz niepoznawczych (ang. *non-cognitive*), pozwalających na lepszą selekcję przyszłych studentów, którzy nie tylko osiągną sukces zawodowy czy naukowy w pielęgniarstwie, lecz także charakteryzują się wysokim stopniem motywacji do podjęcia studiów na tym kierunku, co zmniejsza ryzyko nieukończenia nauki [50-52].

W obliczu narastającego problemu niedoboru kandydatów na studia kształcące w zakresie pielęgniarstwa i wobec wzrostu niepowodzeń wśród studiujących ten kierunek zwiększenie trafności naboru wydaje się konieczne. Lepsze dopasowanie kryteriów kwalifikacyjnych można osiągnąć między innymi przez posługiwanie się narzędziami oceny zdolności i predyspozycji wykraczającymi poza tradycyjny pomiar cech poznawczych kandydata. W przyszłości, jeśli takie rozwiązania będą stosowane, również będą wymagały ewaluacji w celu oszacowania parametrów jakościowych, podobnie jak w przypadku obecnie używanych metod selekcji.

2. OSZACOWANIE TRAFNOŚCI PROGNOSTYCZNEJ KRYTERIÓW NABORU NA STACJONARNE STUDIA PIERWSZEGO STOPNIA NA KIERUNKU PIELĘGNIARSTWO

Na podstawie uzyskanych wyników analizy korelacyjnej można stwierdzić, że wszystkie trzy kryteria przedmiotowe (wyniki matury z języka polskiego, języka obcego i przedmiotu dodatkowego), będące podstawą punktacji rankingowej, pozostają niezależnie od roku studiów w pozytywnej relacji z wynikami kształcenia studentów wyliczonymi jako średnia ocen. Najwyższą trafnością prognostyczną spośród kryteriów kwalifikacyjnych charakteryzował się przedmiot dodatkowy, co prawdopodobnie ma związek z faktem, że w grupie przyjętych kandydatów dominowały osoby, które wybierały przedmiot premiowany (biologia, chemia, fizyka lub wiedza o społeczeństwie). Przedmioty te, a w szczególności biologia, to bardzo ważna podstawa efektywnego kształcenia kompetencji klinicznych. Wydaje się, że wprowadzenie do systemu rekrutacji przedmiotów preferowanych (mnożnik punktacji rankingowej 1,4) sprzyja selekcjonowaniu kandydatów o szczególnych predyspozycjach wejściowych i dobrych podstawach w zakresie nauk przyrodniczych/ścisłych. Ustalenia te pozostają w zgodzie z wnioskami, jakie poczyniono w zakresie oceny zdolności dyskryminacyjnej poszczególnych kryteriów kwalifikacyjnych. Trzeba jednak zaznaczyć, że w stosunku do wyników analizy korelacyjnej należy zachować pewną ostrożność w zakresie interpretacji.

Dobre wyniki w zakresie trafności, uzyskane w jednowymiarowej analizie korelacji, zostały skonfrontowane z danymi uzyskanymi metodą analizy regresji i kowariancji (ANCOVA). Standardowo w analizach regresji, jako ważne predyktory, włącza się do modelu zmienne socjodemograficzne, które są związane z indywidualnym profilem studenta: płeć, wiek w momencie przyjęcia na studia,

rodzaj ukończonej szkoły średniej czy wykształcenie rodziców [24, 55, 56]. Czynniki te uznano za niezależne predyktory, w związku z tym badanie ich wpływu na osiągnięte wyniki kształcenia w trakcie studiów zostało ocenione w oddzielnym modelu regresji.

Z analizy regresji wynika, że mężczyźni osiągają istotnie lepsze wyniki w toku kształcenia na kierunku pielęgniarstwo niż kobiety, i to zarówno w ocenie krótko-, jak i długoterminowej (odpowiednio $\beta_{\text{stand.}}$ 0,057 i 0,054). Wpływ płci na wyniki kształcenia na kierunku pielęgniarstwo nie jest jednoznaczny [57-59]. Różnice w doniesieniach naukowych mogą mieć swoje podłoże kulturowe. O ile z jednej strony ustalenia pochodzące z badań włoskich [58], australijskich [59] i amerykańskich [57] nie wskazują na wpływ płci, o tyle z drugiej strony te dotyczące studentów pakistańskich dają wynik przeciwny [60]. Ponadto pewnym istotnym ograniczeniem metodologicznym wszystkich wymienionych badań jest niski procentowy udział mężczyzn w całej grupie studiujących pielęgniarstwo. Zastrzeżenie to dotyczy także wyników prezentowanych w niniejszym opracowaniu. W celu jednoznacznego określenia roli płci w ocenie jej wpływu na sukces akademicki konieczne jest przeprowadzenie celowanych badań prospektywnych w tym zakresie.

Drugim obok płci potencjalnym predyktorem sukcesu jest zmienna wiek. Z uzyskanych wyników analizy regresji wynika, że młodsi studenci osiągają wyższą średnią ocen, ale jedynie w ocenie długoterminowej (średnia na koniec studiów, $\beta_{\text{stand.}}$ -0,304). Także powiązana z wiekiem zmienna dotycząca przyjęcia na studia w roku zdania matury okazała się istotnym predyktorem sukcesu. Generalnie z danych literaturowych wynika, że w nauce na studiach pielęgniarstkich starsi studenci osiągają lepsze wyniki niż młodsi [61-64]. Jednakże istnieją w tym zakresie też doniesienia, które nie potwierdzają istotnego wpływu tej zmiennej na osiągnięcia edukacyjne [58, 60, 65, 66].

W porównaniu do doniesień naukowych oceniających wartość predykcyjną wieku jest dostępnych niewiele danych, które odnoszą się do wpływu zmiennej *czas od zdania matury do przyjęcia na studia pielęgniarstkie*. Ali i Naylor w swoim badaniu nie zaobserwowali istotnego wpływu przerwy między podjęciem studiów a zakończeniem poprzedniej szkoły [60]. Autorzy podkreślają, że nie ma znaczenia, czy do studiów zostaną dopuszczeni kandydaci zaraz po ukończeniu szkoły średniej czy też tacy, którzy z jakichś powodów mieli przerwę w edukacji [60]. Wobec powyższych ustaleń można przyjąć, że uzyskane wyniki analizy regresji dotyczące wieku i roku przyjęcia na studia na WUM są zbieżne z wnioskami opublikowanymi w literaturze światowej.

Ostatnią z badanych zmiennych socjodemograficznych mających potencjalny wpływ na sukces akademicki była miejscowość ukończenia szkoły średniej

(Warszawa lub inna). Nie wykazano, żeby kandydaci zamiejscowi mieli przewagę w odniesieniu do uzyskanej średniej ocen, ani w prognozie krótko-, ani długoterminowej. Należy zaznaczyć, że miejsce pochodzenia ma silny związek z czynnikami psychologicznymi i ekonomicznymi, które nie były brane pod uwagę w niniejszym opracowaniu [67]. Dodatkowo dane literaturowe oraz teoretyczne modele sukcesu akademickiego wskazują, że poza podstawowym katalogiem zmiennych socjodemograficznych, takich jak płeć czy wiek, istnieje jeszcze wiele innych potencjalnych predyktorów mających wpływ na osiągnięte wyniki kształcenia na studiach pielęgniarstwie [32].

Zgodnie z prezentowanymi na wstępie wynikami analizy korelacyjnej, uzupełnionymi danymi z oceny metodą regresji, można wnioskować, że dwa z trzech przedmiotowych kryteriów kwalifikacyjnych były istotnymi czynnikami warunkującym odniesienie sukcesu podczas studiów. W tym zakresie można więc mówić o potwierdzonej trafności kryterialnej (prognostycznej) stosowanych na WUM zasad kwalifikacji kandydatów na stacjonarne studia pielęgniarstwie pierwszego stopnia. Oznacza to, że w większości przypadków punktacja rankingowa uzyskiwana przez kandydatów podczas rekrutacji odzwierciedlała rzeczywiste właściwości i cechy zdających.

Odmienne wyniki trafności prognostycznej uzyskano w przypadku kryteriów dodatkowych. Jednym z istotnych elementów systemu kwalifikacji kandydatów na WUM jest stosowanie premii punktowej za wybór przedmiotów dodatkowych, takich jak biologia, chemia, fizyka i wiedza o społeczeństwie. Założeniem takiego sposobu prowadzenia naboru jest promowanie tych kandydatów, którzy charakteryzują się odpowiednim stopniem przygotowania akademickiego w zakresie między innymi nauk biologicznych i ścisłych. Jednakże, jak wynika z przeprowadzonych badań regresyjnych, uzupełnionych wnioskami z analizy kowariancji, studenci, którzy wybierali przedmiot premiowany, nie tylko nie mieli lepszych wyników niż pozostali, ale ich osiągnięcia edukacyjne były systematycznie gorsze w prognozie krótkoterminowej (średnia po pierwszym roku studiów, $\beta_{\text{stand.}} -0,335$) i długoterminowej (średnia z toku studiów, $\beta_{\text{stand.}} -0,344$).

Drugim dodatkowym kryterium kwalifikacyjnym, uzupełniającym system naboru kandydatów, była możliwość uzyskania punktów rankingowych za ukończenie kursu kwalifikacyjnego pierwszej pomocy czy też posiadanie dyplomu zawodowego. W przypadku oceny determinantów sukcesu wykazano, że kandydaci przyjęci na studia na podstawie dodatkowej punktacji uzyskiwali istotnie niższe wartości średniej ocen po pierwszym roku nauki i pod koniec studiów (odpowiednio $\beta_{\text{stand.}} -0,123$ i $-0,140$). Reasumując, można stwierdzić, że stosowanie w ramach kwalifikacji kandydatów dodatkowej punktacji nie daje oczekiwanych wyników w zakresie podniesienia jakości naboru. Premiowanie punktacją rankingową

dotychczasowych kwalifikacji (50, 100 lub 150 pkt) nie wpływa pozytywnie na osiągnięte w toku kształcenia sukcesy akademickie studenta mierzone średnią ocen, zarówno w prognozach krótko-, jak i długoterminowych.

Podsumowując ocenę jakości procesu naboru kandydatów na studia pielęgniarstwa na WUM, należy podkreślić, że o ile trzy przedmiotowe kryteria kwalifikacyjne spełniają swoje zadanie w zakresie trafności, o tyle dwa pozostałe (nieobligatoryjne) dodatkowe czynniki selekcyjne obniżają wiarygodność procesu rekrutacji. System naboru nie spełnia swojej funkcji, jeśli pewna grupa studiujących nie jest w stanie sprostać wymaganiom ze względu na brak odpowiednich cech i predyspozycji, które nie zostały prawidłowo ocenione i zweryfikowane podczas przyjmowania na studia. Dlatego też tak ważne jest monitorowanie i stała kontrola przyjętych założeń do strategii rekrutacyjnej, aby uniknąć wymiernych strat ponoszonych w przyszłości przez uczelnię.

3. ANALIZA SPÓJNOŚCI SYSTEMU OCENIANIA OSIĄGNIĘĆ EDUKACYJNYCH STUDENTÓW ORAZ JEGO ZGODNOŚCI Z ZAŁOŻENIAMI ZAWARTYMI W KIERUNKOWYCH STANDARDACH KSZTAŁCENIA

Najczęstszym źródłem niezgodności w odniesieniu do wyników pomiaru (oceny) jest nadmierna surowość/łagodność egzaminatora, a także skłonność do ocen skrajnych lub nadmiernego uśredniania [13]. Jak wskazują wyniki analizy dotyczącej ocen z poszczególnych przedmiotów, w przypadku anatomii i farmakologii studenci uzyskiwali wyraźnie niższe noty w porównaniu z pozostałymi przedmiotami (surowy egzaminator), podczas gdy dla języka obcego oceny były skrajnie wysokie (nadmiernie łagodny egzaminator). Równocześnie w przypadku anatomii zaobserwowano szeroki zakres zmienności wyników, co w naturalny sposób odzwierciedla zróżnicowanie osiągnięć w badanej grupie studentów. Skrajne wartości nie muszą oczywiście świadczyć o braku obiektywizmu punktowania, ale mogą wynikać z faktu, że dany przedmiot wymaga złożonych i trudnych do opanowania kompetencji (jak w przypadku farmakologii – bardzo niska średnia ocen studentów) lub wręcz przeciwnie: osiągnięcie efektów kształcenia w danym zakresie jest dla studentów stosunkowo łatwe (wysoka średnia ocen, jak w przypadku języka obcego). Zróżnicowany dla poszczególnych przedmiotów poziom oczekiwań wobec studentów może w pewnym stopniu uzasadniać wyniki analizy osiągnięć studentów, które wskazują na istotne statystycznie odchylenie w spójności oceniania. Jednakże zgodność oceny w zakresie tego samego przedmiotu była różna w kolejnych latach. Należałoby oczekiwać, że niezależnie od rocznika dobry student w porównaniu z mniej zdolnym będzie miał istotnie wyższe wyniki ze wszystkich przedmiotów.

Rzetelność pomiaru oznacza powtarzalność uzyskanych wyników w określonych warunkach. Jak pokazują przedstawione wyniki badania rzetelności za pomocą współczynnika alfa-Cronbacha, system oceniania osiągnięć studentów charakteryzował się dostatecznym poziomem rzetelności (spełnione kryterium Juma Nunnally'ego przy $\alpha > 0,70$) [46]. Należy podkreślić, że zły dobór metod egzaminacyjnych, a w szczególności nieprawidłowa konstrukcja zadań sprawdzających, które są podstawą oceniania, przyczynia się do spadku rzetelności pomiaru. Student może nie mieć sposobności pokazania pełni swoich osiągnięć z danej dziedziny, jeżeli egzamin znacząco zawęży zakres treści służących ocenie danych efektów kształcenia. Niska rzetelność systemu oceniania może wystąpić szczególnie w przypadkach z bardzo dużym stopniem zróżnicowania ocenianych efektów, które w większości sytuacji, szczególnie w kształceniu ogólnym, dotyczy bardzo odległych cech i właściwości zdających. Inną przyczyną niskiej wartości współczynnika alfa ($< 0,70$) może być duży udział błędów losowych w wynikach danego pomiaru. Przypadkowe wahania wyników oceniania, opisane w klasycznej teorii testu egzaminacyjnego, mogą obniżać wartość współczynnika alfa [68]. Przy wartości alfa $< 0,7$ błędy losowe stanowią więcej niż 30% zmienności uzyskanych wyników, a pomiar w takich warunkach według Joya Guilforda może być stosowany właściwie jedynie przy porównywaniu międzygrupowym, a nie do różnicowania indywidualnego [69].

Analizując wartości indeksów dyskryminacyjnych dla poszczególnych przedmiotów z czterech grup efektów kształcenia, można stwierdzić, że wyniki z zakresu fizjologii (indeks 0,30) i zdrowia publicznego (indeks 0,34) odbiegają od ogólnej oceny osiągnięć studentów. Szczególnie wysokie wartości mocy różnicującej zanotowano dla anatomii (indeks 0,48; blok A), języka obcego (indeks 0,46; blok B), podstawowej opieki zdrowotnej (indeks 0,45; blok C) oraz chirurgii (indeks 0,48; blok D). Wyniki egzaminów z tych czterech przedmiotów stanowią punkt centralny w systemie oceny osiągnięć studentów w poszczególnych grupach efektów kształcenia. Można założyć, że zasady oceniania studentów przyjęte dla tych przedmiotów mogą być traktowane jako rozwiązanie modelowe i powinny stanowić podstawę poprawy parametrów różnicujących dla pozostałych przedmiotów z danej grupy efektów kształcenia.

Zgodnie z kryterium Kaisera struktura funkcjonującego na WUM systemu oceny osiągnięć studentów przyjęła postać trójelementową, co nie jest zgodne z założeniami teoretycznymi. Ponadto wyniki eksploracyjnej analizy czynnikowej wskazują, że rozłokowanie ocen studenckich dla 16 przedmiotów w strukturze skali pomiaru osiągnięć wyjaśnia mniej niż połowę obserwowanej ogólnej zmienności (42,0%). Poszczególne składowe nie były homogenne w zakresie poszczególnych grup przedmiotów. Oprócz tego wyniki analizy czynnikowej są w dużym stopniu spójne z wynikami oceny zgodności punktowania oraz pomiarami rzetelności

za pomocą współczynnika alfa-Cronbacha, które wykazały, że oceny uzyskane z przedmiotu fizjologia są odmienną miarą cech i właściwości zdających w stosunku do wyników egzaminowania z pozostałych przedmiotów. Także analiza jednowymiarowości oceniania dla każdej z grup przedmiotów wykazała, że w dwóch przypadkach skala oceny osiągnięć studentów nie jest jednorodna. Dotyczyło to grup przedmiotów z bloków A i D. Podsumowując wyniki dotyczące szacowania trafności teoretycznej oraz jednowymiarowości, można stwierdzić, że oceny z poszczególnych przedmiotów nie odzwierciedlają podziału grupowego efektów kształcenia zaproponowanego w standardach kształcenia na kierunku pielęgniarstwo.

Drugim analizowanym aspektem trafności była ocena zgodności kryterialnej wyników uzyskiwanych przez studentów z poszczególnych przedmiotów egzaminacyjnych. Opierając się na danych z analizy korelacyjnej, można stwierdzić, że podobnie jak w przypadku trafności teoretycznej, niespełnione są założenia dla trafności kryterialnej we wszystkich czterech grupach przedmiotów. Wyniki analizy interkorelacji wskazują, że dla wszystkich przedmiotów, z wyjątkiem fizjologii, istnieją pozytywne zależności między wynikami uzyskanymi przez studentów z kolejnych egzaminów. Jednakże siła korelacji dla poszczególnych par przedmiotów była zróżnicowana, a średnia współczynnika korelacji r -Spearmana (r_s) oscylowała wokół wartości 0,20 i była wyraźnie niższa tylko w przypadku wyników dla fizjologii (średnia $r_s = 0,12$). W obrębie tej samej domeny utworzonej z przedmiotów należących do wspólnej grupy efektów kształcenia szczególnie wysokie wartości korelacji zanotowano dla par przedmiotów: anatomia i farmakologia ($r_s = 0,28$), podstawowa opieka zdrowotna i podstawy pielęgniarstwa ($r_s = 0,27$) oraz psychiatria i ginekologia ($r_s = 0,33$).

W zakresie analizy korelacji dla zależności: wyniki kształcenia z nauk podstawowych *versus* z zakresu opieki specjalistycznej szczególnie korzystne relacje zaobserwowano w przypadku par: farmakologia *versus* chirurgia ($r_s = 0,31$) oraz parazytologia *versus* choroby wewnętrzne ($r_s = 0,31$). Wyniki te w doskonały sposób prezentują pozytywną zależność między oceną z pierwszego przedmiotu (zajęcia w pierwszym roku nauki) a wynikami studenta osiągniętymi w ramach kształcenia z przedmiotu klinicznego (drugi i trzeci rok nauki). Korelacja taka dowodzi silnego związku między osiągniętymi efektami kształcenia dla obu przedmiotów. Z drugiej strony zaobserwowano, że dla takiego przedmiotu przedklinicznego jak fizjologia istnieje nikła korelacja z ocenami z zakresu neurologii ($r_s = 0,04$). Poprawa parametrów spójności oceniania wymaga szczegółowej analizy jakościowej stosowanych metod pomiaru edukacyjnego dla wszystkich przedmiotów egzaminacyjnych.

Stosunkowo wysoka rzetelność systemu oceniania przyczynia się do względnie dużej dokładności (precyzji) oceny wiedzy i umiejętności zdających, a co za tym idzie – dobrego różnicowania indywidualnego. Jednakże spore zastrzeżenia budzą wyniki analizy trafności. W niektórych przypadkach można uznać, że wyniki uzyskiwane przez studentów nie odzwierciedlają rzeczywistych cech zdających (*vide* fizjologia). Należy jednak, w pogłębionej analizie programu kształcenia, dokonać krytycznego osądu, na ile taki przedmiot rzeczywiście reprezentuje zakres kompetencji istotnych dla późniejszego kształcenia klinicznego. Sytuacja, w której dość dobra rzetelność idzie w parze z niedostateczną trafnością, skutkuje obniżeniem wiarygodności całego systemu oceniania kompetencji. Ma to szczególne znaczenie dla oceny tych studentów, których średnia ocen z toku studiów lokuje się w pobliżu dolnej granicy skali (średnia ocen $\approx 3,0$). Ze względu na fakt, że są to studenci najslabsi, istnieje duże ryzyko niedostatecznego poziomu kompetencji reprezentowanych przez tę grupę absolwentów pielęgniarstwa.

Przedstawiona powyżej analiza trafności pomiaru edukacyjnego ma za zadanie zapobiegać nadużyciom w interpretacji wyników pomiaru [70]. Jeżeli student uzyskał wysoką średnią ocen z toku studiów, to wartość takiej oceny jest istotna tylko wtedy, gdy odzwierciedla jego rzeczywiste osiągnięcia, w szczególności w odniesieniu do wymagań programowych. Dlatego też jednym z ważnych aspektów analizy jakości pomiaru jest szacowanie trafności prognostycznej, która odnosi się do oceny, w jakim stopniu dane wyniki kształcenia mogą służyć do przewidywania przyszłych losów studentów, np. statusu zawodowego absolwentów. W analizowanym przypadku określenie trafności prognostycznej z oczywistych względów nie jest do końca możliwe. Nie dysponujemy bowiem szczegółowymi danymi o przyszłych losach absolwentów w zakresie ich aktywności zawodowej, po zakończeniu studiów. Jednakże w pewnym stopniu ocenę trafności prognostycznej systemu oceniania na studiach pierwszego stopnia przeprowadzono, opierając się na wynikach kształcenia absolwentów WUM, którzy podjęli na tej samej uczelni studia drugiego stopnia. Wyniki tej analizy zostaną zaprezentowane w części dotyczącej szacowania edukacyjnej wartości dodanej.

4. WYTYPOWANIE ISTOTNYCH DETERMINANTÓW SUKCESU/PORAŻKI AKADEMICKIEJ STUDENTÓW, KTÓRZY PODJĘLI KSZTAŁCENIE NA KIERUNKU PIELĘGNIARSTWO

Istotnym elementem walidacji założeń polityki rekrutacyjnej uczelni jest identyfikowanie czynników predykcyjnych, które trafnie oceniałyby prawdopodobieństwo odniesienia sukcesu przez kandydata w trakcie nauki, a także tych czynników, które przyczyniają się do wzrostu ryzyka niepowodzenia podczas studiów. Prowadzenie badań identyfikujących czynniki wpływające na ryzyko porażki edukacyjnej ma na celu określenie profilu studenta, który jest

szczególnie narażony na problemy z terminowym ukończeniem studiów, lub też obciążonego wysokim ryzykiem rezygnacji w trakcie kształcenia. Jeśli będziemy w stanie wytypować studentów z grupy ryzyka, to możliwe będzie podjęcie takich działań wspierających, które skutecznie zmniejszą odsetek niepowodzeń.

W zasobach piśmiennictwa jest dostępnych kilka prac opisujących relacje między zmiennymi demograficznymi a późniejszymi wynikami osiąganymi podczas studiów na kierunku pielęgniarstwo [71-74]. Jako ważne predyktory sukcesu wymieniane są: płeć, wiek, rodzaj ukończonej szkoły średniej czy wykształcenie rodziców. Wobec powyższego można zadać pytanie, czy w grupie czynników ryzyka niepowodzenia mogą się znaleźć także zmienne socjodemograficzne.

Testowany model logistyczny wskazuje na istotne znaczenie miejsca ukończenia szkoły średniej. Studenci spoza Warszawy byli obciążeni o ponad 35% mniejszym ryzykiem niepowodzenia w porównaniu ze studentami miejscowymi (OR = 1,358; p = 0,023). Obserwację tę można tłumaczyć faktem, że w grupie studentów zamiejscowych są osoby, których wybór uczelni i kierunku był decyzją dobrze przemyślaną, ze względu na konieczność zorganizowania pobytu poza dotychczasowym miejscem zamieszkania. Grupa ta powinna się więc charakteryzować większą motywacją, determinacją i zaangażowaniem w proces kształcenia. Jednakże z drugiej strony w przypadku studentów zamiejscowych można byłoby się spodziewać zwiększonego ryzyka niepowodzenia ze względu na wpływ czynników ekonomicznych, w związku z wyższym kosztem studiowania poza dotychczasowym miejscem zamieszkania. Ponadto studenci zamiejscowi mogą mieć większe problemy z adaptacją i integracją społeczną, szczególnie w przypadku tych osób, które przyjechały do Warszawy z małych miast i wsi.

Wyniki analizy modelu regresji logistycznej dają również podstawę do wnioskowania o znaczącej roli płci studenta jako istotnego predyktora niepowodzenia. Wykazano, że mężczyźni mają ponad 2,5-krotnie większe szanse na niepowodzenie podczas studiów niż kobiety (OR = 2,641; p < 0,001). Badania jakościowe prowadzone metodą wywiadu dowodzą, że w przypadku występowania silnej dominacji kobiet w środowisku akademickim oraz zawodowym, z czym mamy do czynienia w przypadku pielęgniarstwa, prawdopodobieństwo podjęcia przez mężczyznę decyzji o dobrowolnej rezygnacji z dalszej nauki lub pracy w zawodzie jest bardzo wysokie [75].

Ani wiek studenta w momencie rozpoczęcia studiów, ani przyjęcie na studia w roku zdania matury nie miało istotnego wpływu na ryzyko porażki. Generalnie w zakresie wpływu wieku studenta na status ukończenia studiów doniesienia różnych autorów wskazują, że wśród młodych osób notuje się istotnie większe ryzyko porażki niż w pozostałej grupie studentów [72, 76-80].

Jak wskazują uzyskane wyniki analizy kryteriów kwalifikacyjnych, punktacja rankingowa wyliczana dla języka polskiego, języka obcego i przedmiotu dodatkowego nie wpływała w istotny sposób na status ukończenia studiów. Oznacza to, że w grupie osób, które nie ukończyły nauki, nie ma studentów z czołówki listy rankingowej (nie zanotowano istotności statystycznej dla $OR > 1$). Jednakże dowodzi to również tego, że stosowany system kwalifikacji nie selekcjonuje kandydatów w taki sposób, iż niższe pozycje w rankingu zajmują osoby, dla których ryzyko porażki byłoby większe w stosunku do studentów, którzy uzyskali wyższą punktację rankingową (nie zanotowano istotności statystycznej dla $OR < 1$). Oczekuje się, że lepiej przygotowani kandydaci na studia powinni być w niewielkim stopniu obarczeni ryzykiem nieukończenia studiów. Jednakże dostępne dane literaturowe wskazują, że są sytuacje, w których studenci o dobrym i bardzo dobrym przygotowaniu akademickim nie kończą nauki na wybranym przez siebie kierunku. Mulholland i wsp. [71] stwierdzili, że kandydaci o wyższych kwalifikacjach wejściowych bardzo często podejmują świadomą decyzję o rezygnacji z dalszego studiowania na kierunku pielęgniarstwo, zwykle z powodu zmiany kierunku studiów i/lub uczelni. Status ukończenia studiów nie jest więc w tych przypadkach związany z niskim poziomem osiągnięć edukacyjnych studenta.

W zgodzie z powyższymi obserwacjami pozostają wyniki dotyczące negatywnego wpływu zmiennej przedmiot premiowany na ryzyko nieukończenia studiów pielęgniarzkich na WUM ($OR = 2,024$; $p = 0,005$). Na podstawie analizy predykcyjnej można wnioskować, że studenci wybierający biologię, chemię, fizykę lub wiedzę o społeczeństwie na egzaminie maturalnym są obarczeni znacznie większym ryzykiem nieukończenia studiów niż pozostali studenci kierunku. Jak wskazują wyniki badań dotyczące studentów pielęgniarstwa prowadzone w latach 90. w Irlandii Północnej [81], przyczyn wysokiego wskaźnika rotacji wśród studentów z dobrymi i bardzo dobrymi kwalifikacjami wejściowymi należy upatrywać głównie w motywacjach osobistych, a nie w słabych wynikach nauczania. Powyższe przesłanki sugerują więc, że w badaniach predykcyjnych dotyczących oceny czynników ryzyka nieukończenia studiów należy uwzględnić szeroki katalog przyczyn usunięcia osoby z listy studentów.

STUDIA DRUGIEGO STOPNIA

1. OCENA PRZEBIEGU PROCESU SELEKCJI KANDYDATÓW NA STUDIA STACJONARNE Z UWZGLĘDNIENIEM ZMIAN, JAKIE ZACHODZIŁY W NABORZE W LATACH 2005–2015

W Polsce każda uczelnia stosuje własne metody naboru na studia drugiego stopnia na kierunku pielęgniarstwo. W części z nich obowiązuje niestandardyzowany testowy egzamin wstępny, obejmujący swoim zakresem tematykę studiów pierwszego stopnia. Nie ma więc możliwości porównania jakości tak prowadzonego naboru, ze względu na brak jednolitych krajowych kryteriów, jak ma to miejsce w przypadku przyjęć na studia pierwszego stopnia, gdzie obowiązuje ranking oparty na wynikach przedmiotowych egzaminów maturalnych.

Analiza dyskryminacyjna dostarczyła informacji w odniesieniu do tego, w jakim stopniu poszczególne kryteria kwalifikacyjne klasyfikują kandydatów do jednej z dwóch kategorii: przyjęty/nieprzyjęty. Na podstawie otrzymanych wyników można przyjąć, że jedynie punktacja uzyskana z testowego egzaminu wstępnego charakteryzuje się dostateczną zdolnością selekcyjną w wyborze kandydatów na kierunek pielęgniarstwo studia drugiego stopnia na WUM. Analizując wartości AUC, przebieg krzywych ROC oraz wielkość indeksów Youdena w kolejnych latach naboru, można stwierdzić, że wynik egzaminu testowego miał zasadniczy wpływ na podjęcie decyzji o przyjęciu/nieprzyjęciu na studia. Kryterium to odznaczało się równocześnie wysoką czułością i swoistością, co zapewniło, w porównaniu ze średnią ocen (drugie kryterium), powtarzalną i sprawiedliwą ocenę kompetencji wszystkich kandydatów przystępujących do kwalifikacji na studia. Także ustalenia, które wynikają z analizy metodą regresji logistycznej, wskazują, że to punktacja kwalifikacyjna za egzamin testowy stanowiła główny predyktor sukcesu podczas naboru na studia na WUM.

Jednakowe traktowanie wszystkich zdających w kolejnych latach oznacza niezależną ocenę ich osiągnięć, bez względu na to, jakie wyniki osiągnęli w poprzednich cyklach kształcenia. Jest to szczególnie zaleta egzaminu wstępnego w porównaniu z kryterium średniej ocen, gdzie mamy do czynienia z niejednolitym systemem oceniania w uczelniach prowadzących studia pierwszego stopnia. Problem skuteczności selekcji opartej na wartości średniej ocen jest dostrzegany przez wielu badaczy, którzy zwracają uwagę na zróżnicowanie systemów oceniania i niską rzetelność takiego pomiaru selekcyjnego [5]. Wybór obiektywnych kryteriów stosowanych w procesie kwalifikacji kandydatów jest niezbędny do przeprowadzenia wiarygodnej selekcji. Obawy co do subiektywności oceny nie tylko dotyczą kryterium średniej ocen, lecz także, jak podają niektóre doniesienia naukowe, mogą się odnosić do takich metod, jak kierowana rozmowa kwalifikacyjna czy esej refleksyjny [82, 83].

Kolejnym obok obiektywizmu ważnym aspektem oceny jakości procesu selekcji kandydatów jest zachowanie w miarę wysokiej zgodności oceny kompetencji między poszczególnymi kryteriami. Należy oczekiwać, że niezależnie od rocznika w puli przyjętych na studia będą dominowali kandydaci, którzy uzyskali wysokie wyniki dla każdego z kryteriów. Obecne zasady naboru na WUM przewidują zróżnicowanie wag punktacji rankingowej dla poszczególnych kryteriów. Nie ustalono jednak minimalnej punktacji dla każdego kryterium z osobna. Jak wynika z przebiegu poszczególnych krzywych ROC, w niektórych zakresach punktacji mamy do czynienia z wysokimi wynikami z egzaminu wstępnego uzyskanymi przez kandydata, przy równocześnie niskiej wartości średniej ocen z toku studiów pierwszego stopnia, i *vice versa*. Rozmijanie się punktacji za te dwa kryteria sugeruje, że w przypadku niektórych kandydatów oba pomiary kompetencji nie są rzetelne i trafne. Mając na uwadze przedstawione powyżej zastrzeżenia co do obiektywności oceny za pomocą średniej ocen, można zakładać, że właśnie to kryterium odpowiada za brak obserwowanej zgodności obu pomiarów.

Z przeprowadzonej analizy dyskryminacyjnej wynika, że egzamin testowy był najważniejszym kryterium kwalifikacyjnym. W związku z tym, aby utrzymać wiarygodność procesu selekcji kandydatów, kluczowym elementem jest zapewnienie wysokiego poziomu jakości pomiaru dokonywanego tym narzędziem. Ponieważ stosowany na WUM testowy egzamin wstępny jest narzędziem niestandardyzowanym, o nieznanymi właściwościach psychometrycznych, to szczególnie ważne jest retrospektywne oszacowanie jego jakości, zwłaszcza w zakresie trafności i rzetelności.

Ważną właściwością dobrych narzędzi selekcyjnych jest ich rzetelność, rozumiana jako powtarzalność uzyskanych wyników (miara precyzji pomiaru) [84]. Niedostateczna rzetelność stosowanej procedury kwalifikacyjnej przyczynia się do niskiego stopnia zaufania. Jak pokazują uzyskane wyniki badania rzetelności za pomocą współczynnika alfa-Cronbacha, spójność wewnętrzna w przypadku wszystkich analizowanych roczników nie była zadowalająca. Mimo że ocena rzetelności uzyskana za pomocą wyliczenia współczynnika Guttmana wskazuje na nieco wyższą zgodność pomiaru, niż to miało miejsce w przypadku obliczeń otrzymanych metodą Cronbacha, to nadal nie były to wyniki zadowalające. Szacując rzetelność poszczególnych edycji egzaminu wstępnego różnymi metodami analitycznymi, należy mieć na uwadze, że każda ocena rzetelności uwzględnia inne źródło błędów, stąd pewne rozbieżności między uzyskanymi wynikami [84].

Ponieważ istotnym czynnikiem, który ma wpływ na rzetelność pomiaru za pomocą serii pytań wielokrotnego wyboru, jest prawidłowa konstrukcja testu, to ważna jest ocena jakości poszczególnych zadań tworzących pulę egzaminacyjną [85]. Porównując poszczególne edycje egzaminu, zaobserwowano

zróznicowany poziom ogólnej trudności, od stosunkowo łatwego testu w roku 2015/16 do średnio trudnego w 2010/11 (indeks łatwości odpowiednio 0,71 i 0,49). Także pod względem mocy różnicującej kolejne edycje egzaminu nie były jednorodne. Duża liczba pytań łatwych i aprobujących, czyli takich, które wymagają od zdającego jedynie potwierdzenia poprawności podanej informacji (szczególnie z opcjami do wyboru, typu „wszystkie z powyższych” lub „żaden z powyższych”), jest intelektualnie mniej wymagająca i przyczynia się do uzyskania przez kandydatów przeciętnie wyższych wyników z takiego egzaminu [86, 87]. Ważne jest więc dobranie takich pytań egzaminacyjnych, które wykazują znaczną moc różnicującą. Tego rodzaju pytania pozwalają na odróżnianie kandydatów, którzy dla danej mierzonej wartości uzyskują wyniki istotnie różne, co pozwala na dobrą selekcję [88]. W literaturze pytania takie są określane jako zadania rozrywające, ponieważ powodują spłaszczenie rozkładu wyników (kurtoza mniejsza od zera), a to znacznie ułatwia dyskryminację i ustalenie punktu odcięcia: przyjęty/nieprzyjęty [88]. Ta ważna cecha egzaminów selekcyjnych, mierzona za pomocą średniej wartości dla mocy różnicującej wszystkich pytań z puli testowej, była wyraźnie niższa w przypadku edycji egzaminu w latach 2009/10 i 2011/12 ($\approx 0,13$). Dość wysokie były natomiast wyniki średniej mocy różnicującej dla dwóch ostatnich egzaminów wstępnych: z 2014/15 i 2015/16 roku ($\approx 0,24$).

Jednym z dostępnych sposobów poprawy parametrów jakościowych zadań testowych jest zmiana przypadającej na każde pytanie liczby opcji do wyboru. Mimo że specjaliści zajmujący się analizami psychometrycznymi od wielu lat podkreślają, że trzy opcje to optymalna liczba pozycji w testach egzaminacyjnych [89, 90], to nadal duża grupa nauczycieli akademickich stosuje pytania cztero- i pięcioopcyjne [91]. Michael Rodriguez wykazał w metaanalizie, że efekt łatwiejszego zgadywania w przypadku mniejszej liczby opcji jest powszechnie przeceniany [90]. Liczne wyniki badań wykazały, że zmniejszenie liczby opcji z czterech/pięciu do trzech skutkowało wzrostem wyników w granicach 1,0–1,2% [89–92]. Wobec powyższych argumentów wydaje się, że decyzja organizatorów testowego egzaminu wstępnego dotycząca zwiększenia liczby opcji w pytaniach testowych nie była do końca racjonalna, co potwierdziły wyniki retrospektywnej analizy danych.

Tylko trzy edycje egzaminu wstępnego były oparte na testach zbudowanych z pytań zawierających cztery opcje (2009/10, 2010/2011 i 2012/13), pozostałe cztery roczniki zdawały test złożony z pytań pięcioopcyjnych. Porównując średnie poziomy łatwości dla tych dwóch grup egzaminów, nie wykazano istotnych statystycznie różnic. Zanotowano jednak większą zdolność dyskryminacyjną egzaminów złożonych z pytań zawierających pięć opcji w porównaniu z tymi opartymi na czterech opcjach. Generalnie powyższa obserwacja jest zgodna z doniesieniami naukowymi w tym zakresie. Bardzo obszerna, licząca 56 badań

metaanaliza przygotowana przez Rodrigueza pokazała, że w przypadku dobrze opracowanych pytań testowych wystarczające jest użycie trzech opcji [90]. Powyższe wnioski potwierdzają także przeglądy systematyczne opublikowane przez Haladyna i wsp. [93] oraz Vyasa i Supe'a [89]. Jak podają autorzy, pytania trzyopcyjne mają podobną skuteczność jak pytania czteroopcyjne. Ważną zaletą pytań z mniejszą liczbą opcji jest łatwiejsza konstrukcja, ponieważ twórcy zadań mogą w znacznie większym stopniu zaproponować dobrej jakości dystraktory, które będą efektywniej różnicowały zdających.

Z opublikowanych wyników badań można także wywnioskować, że testy zbudowane z pytań z mniejszą liczbą opcji charakteryzują się lepszą rzetelnością [90, 94]. Przedstawione wcześniej wyniki, które wskazują na niedostateczny poziom rzetelności egzaminów, skłaniają do krytycznego przeglądu stosowanych pytań egzaminacyjnych, polegającego na eliminacji elementów słabo różnicujących, zwiększeniu spójności dziedzinowej zestawu oraz wydłużeniu skali pomiarowej. Szczególnie ostatni z wymienionych sposobów może być dobrą metodą podnoszenia poziomu rzetelności pomiaru. Większa liczba pytań w puli egzaminacyjnej nie musi w tym przypadku oznaczać wydłużenia czasu egzaminowania. Jak bowiem wykazały wyniki metaanalizy opublikowanej przez Aamodta i McShane'a [92], w przypadku 100-pytaniowego egzaminu studenci są w stanie rozwiązać średnio o 12,4 więcej zadań przygotowanych w formie trzech opcji w porównaniu z testem składającym się z pytań czteroopcyjnych. Autorzy tej samej metaanalizy podają również, że stworzenie 100-pytaniowego testu trzyopcyjnego pozwala zaoszczędzić nawet 16 godzin pracy autora takiego egzaminu w porównaniu z przygotowaniem wersji opartej na czterech opcjach [92]. Dlatego też mniej czasochłonne i dużo łatwiejsze w opracowaniu pytania z mniejszą liczbą opcji (trzy zamiast czterech/pięciu) mogą być dobrym rozwiązaniem w przypadku konieczności znacznego zwiększenia rzetelności testu przez wydłużenie skali pomiarowej.

Przeprowadzona przez Autora analiza retrospektywna pozwala na opracowanie dobrego planu testu oraz przygotowanie bogatej bazy pytań egzaminacyjnych z różnych domen wiedzy i umiejętności według ścisłej taksonomii. Rozpoczęty w roku 2014/15 proces budowania bazy danych oraz generowania zestawu pytań testowych za pomocą oprogramowania Bank Zadań Testowych[©] pozwolił na znaczne zwiększenie rzetelności egzaminu oraz poprawę parametrów różnicujących w ostatnich dwóch latach. Działania te powinny zapewnić w przyszłości osiągnięcie wysokiego poziomu rzetelności oceny kompetencji wejściowych kandydatów ubiegających się o możliwość nauki na studiach drugiego stopnia na kierunku pielęgniarstwo na WUM.

2. OSZACOWANIE TRAFNOŚCI PROGNOSTYCZNEJ KRYTERIÓW NABORU NA STACJONARNE STUDIA DRUGIEGO STOPNIA NA KIERUNKU PIELEŃNIARSTWO

Z przeprowadzonej analizy predykcijnej wynika, że oba obligatoryjne kryteria kwalifikacyjne (wynik testu oraz średnia ocen) spełniły swoją funkcję w zakresie selekcji najlepszych kandydatów. Wykazano, że im wyższą średnią ocen uzyskał student na studiach pierwszego stopnia, tym lepsze były jego osiągnięcia podczas kształcenia na studiach drugiego stopnia. Zdolność predykcijną tego kryterium potwierdzają ustalenia Thada Wilsona zebrane w przeglądzie literatury dotyczącym selekcji kandydatów na studia pielęgniarstwa [95]. Mimo pozytywnych wyników analiz należy zaznaczyć, że stosowanie kryterium średniej ocen w procesie naboru na studia nie jest pozbawione pewnych wad. Najczęściej wskazuje się, że systemy oceniania w uczelniach prowadzących studia pierwszego stopnia są bardzo zróżnicowane, szczególnie w zakresie stosowania obiektywnych narzędzi oceny wiedzy i umiejętności studentów. Z jednej strony niejednorodność systemów oceniania w różnych placówkach edukacyjnych, w szczególności w zakresie stosowania standaryzowanych metod oceny, może się przyczyniać do niskiego stopnia rzetelności procesu kwalifikacji kandydatów z zastosowaniem tego kryterium [4]. Z drugiej jednak strony należy pamiętać, że każda uczelnia powinna uzyskać pozytywną akredytację do prowadzenia kształcenia studentów pielęgniarstwa, co jest potwierdzeniem spełnienia wymagań stawianych przez Krajową Radę Akredytacyjną Szkół Pielęgniarek i Położnych. Ponadto jednostka prowadząca studia na tym kierunku powinna zapewnić, że ocena osiągniętych efektów kształcenia będzie spełniała kryteria dobrego pomiaru dydaktycznego, do czego zobowiązują ją odpowiednie zapisy w standardach kształcenia [96]. Po spełnieniu powyższych rekomendacji średnia ocen kandydatów z różnych uczelni może być ze sobą porównywana i stanowić dobre narzędzie różnicujące poziom kompetencji wejściowych kandydata na studia drugiego stopnia.

Drugim analizowanym kryterium kwalifikacyjnym był wynik z testowego egzaminu wstępnego. Wykazano, że sumaryczna punktacja uzyskana z testu jest dobrym wskaźnikiem przyszłych osiągnięć studenta. Jednakże zdolność predykcyjna tego czynnika kwalifikacyjnego była porównywalna z kryterium średniej ocen. Ustalenia te są w pewnym stopniu rozbieżne wobec oczekiwań. Zakładając, że testowy egzamin wstępny jest narzędziem zaprojektowanym specjalnie pod kątem selekcji kandydatów o odpowiednich cechach i predyspozycjach, należałoby oczekiwać, że będzie się on charakteryzował lepszymi parametrami jakościowymi niż kryterium średniej ocen. Wyniki analizy korelacyjnej potwierdzają powyższe założenia, jednak na podstawie ustaleń wynikających z analizy regresji nie można mówić o lepszej czy gorszej trafności średniej ocen czy testu egzaminacyjnego.

Przyczyny niższej spodziewanej trafności prognostycznej egzaminu wstępnego można upatrywać w niepoprawnym przygotowaniu planu testu oraz w niedostatecznej mocy dyskryminacyjnej pytań wchodzących w skład puli zadań egzaminacyjnych. Powyższa teza ma swoje poparcie w opisanych wcześniej wynikach dotyczących jakości tego narzędzia pomiaru kompetencji wejściowych kandydatów. Dodatkowo należy zaznaczyć, że test egzaminacyjny przygotowywany przez nauczycieli akademickich WUM nie jest wystandaryzowany, a więc skala oceny wiedzy i umiejętności zdającego nie spełnia pewnych istotnych parametrów psychometrycznych. Jest to szczególnie ważne w zestawieniu z danymi literaturowymi, w których to walidacji poddaje się specjalnie opracowane trafne, rzetelne i wystandaryzowane narzędzia oceny umiejętności [24, 97].

Trzecim, dodatkowym (nieobligatoryjnym) kryterium kwalifikacyjnym była punktacja przyznawana tylko tym kandydatom, którzy mogli się pochwalić udokumentowaną działalnością naukową w okresie studiów pierwszego stopnia. Wprowadzenie dodatkowej premii punktowej miało na celu przyciągnięcie kandydatów o ponadstandardowych zainteresowaniach naukowych. Absolwenci z dyplomem licencjata, działający wcześniej w kołach naukowych lub będący autorami/współautorami publikacji naukowej albo doniesienia konferencyjnego mogli liczyć na dodatkowe punkty, które były silnym predyktorem sukcesu podczas naboru na studia drugiego stopnia.

Niestety, analizując trafność prognostyczną tego kryterium, wykazano tylko niewielki wpływ tej zmiennej na sukces. Studenci z premią punktową osiągnęli nieco lepsze wyniki w toku studiów, i to zarówno w ocenie krótko-, jak i długoterminowej, aczkolwiek obserwacja ta nie zyskała potwierdzenia w modelowaniu metodą ANCOVA. Przyczyny braku spodziewanych korzyści z przyjęcia osób, które wykazują zainteresowanie badaniami naukowymi, można upatrywać między innymi w fakcie, że miara sukcesu akademickiego, jaką jest średnia ocen z toku studiów, nie odzwierciedla rzeczywistego sukcesu w kształceniu magistrów pielęgniarstwa. Być może uwzględnienie w przyszłości danych dotyczących losów zawodowych absolwentów studiów drugiego stopnia pozwoli na ocenę prognostyczną tego kryterium kwalifikacyjnego w długoterminowej analizie predykcyjnej.

3. ANALIZA SPÓJNOŚCI SYSTEMU OCENIANIA OSIĄGNIĘĆ EDUKACYJNYCH STUDENTÓW ORAZ JEGO ZGODNOŚCI Z ZAŁOŻENIAMI ZAWARTYMI W KIERUNKOWYCH STANDARDACH KSZTAŁCENIA.

Uzyskane wyniki analizy spójności oceniania studentów z poszczególnych przedmiotów dla kolejnych roczników wskazują na wysoki stopień niejednorodności oceniania. Ponadto zaobserwowano statystycznie istotne różnice

w ocenach studentów z danego przedmiotu dla następujących po sobie roczników. Powyższe wyniki wskazują, że niektóre czynniki wpływają na stopień zróżnicowania ocen studentów. Jest mało prawdopodobne, aby istotne różnice w strukturze ocen, mierzone średnią/medianą i odchyleniem standardowym, były spowodowane skokową zmianą ocenianych osiągnięć. Otwarte więc pozostaje pytanie, czy brak powtarzalności w ocenie osiągnięć studentów ma związek ze stronniczością oceny czy też decydujące znaczenie ma tutaj niewłaściwy dobór metod pomiaru dydaktycznego lub ich niska jakość albo zmienność grupy studiujących.

Głównym źródłem niezgodności w uzyskiwanych wynikach pomiaru jest nadmierna surowość lub łagodność egzaminatora, co jest szczególnie widoczne w przypadku przeprowadzania niestandardyzowanych egzaminów ustnych i praktycznych. Także skłonność egzaminatora do stawiania ocen skrajnych może być podstawą do zakwestionowania obiektywizmu punktowania [13].

Z jednej strony, jak wskazują wyniki dotyczące ocen z poszczególnych przedmiotów, z epidemiologii studenci uzyskiwali wyraźnie niższe noty w porównaniu z pozostałymi przedmiotami (surowy egzaminator), a z drugiej – dla podstaw psychoterapii zanotowano wyraźnie wyższą średnią ocen (łagodny egzaminator). Ponadto dla wszystkich analizowanych przedmiotów zaobserwowano szeroki zakres zmienności wyników mierzonych wielkością odchylenia standardowego i współczynnikiem zmienności, co w naturalny sposób odzwierciedla zróżnicowanie osiągnięć w badanej grupie studentów. Obserwowana bardzo niska lub bardzo wysoka średnia ocen dla danego przedmiotu nie musi świadczyć o braku obiektywizmu punktowania. Niskie oceny, które są charakterystyczne dla epidemiologii, mogą mieć związek z faktem, że przedmiot wymaga opanowania złożonych i trudnych umiejętności, a więc w grupie studiujących jest niewiele osób, które są w stanie osiągnąć tego rodzaju efekty kształcenia na wysokim poziomie. Różny stopień oczekiwań w zakresie wiedzy i umiejętności dla poszczególnych przedmiotów może więc być przyczyną obserwowanego odchylenia w międzyprzedmiotowej spójności oceniania.

Znacznie bardziej miarodajne wyniki uzyskano na podstawie oszacowania zgodności ocen otrzymywanych przez studenta z poszczególnych przedmiotów w całym cyklu kształcenia. Należałoby się spodziewać, że niezależnie od rocznika dobry student w porównaniu z mniej zdolnym będzie miał istotnie wyższe wyniki ze wszystkich przedmiotów. Miarą takiej zgodności ocen jest współczynnik Tau Kendalla [98]. Wartość tego współczynnika była bardzo różna dla kolejnych roczników, od stosunkowo niskich (Tau = 0,32) do bardzo wysokich (Tau = 0,66), co świadczy o niezachowaniu we wszystkich przypadkach zasady obiektywizmu punktowania w ocenie studentów w kolejnych rocznikach (łącznie dla badanego okresu Tau = 0,33).

Przedstawione powyżej rozważania dotyczące bezstronności sytuacji pomiarowej oraz obiektywizmu punktowania mogą posłużyć jako ważny argument w dyskusji na temat konieczności wprowadzania w ramach edukacji studentów pielęgniarstwa nowoczesnych metod egzaminowania, zastępujących tradycyjne, niestandardyzowane egzaminy. Warto podkreślić, że zgodnie z obowiązującymi standardami kształcenia na kierunku pielęgniarstwo studia drugiego stopnia, do oceny osiągniętych efektów kształcenia należy stosować zróżnicowane formy, adekwatne do obszarów, których dotyczą te efekty. Zwraca się ponadto uwagę, że w szczególności egzaminy ustne powinny być standaryzowane oraz ukierunkowane na sprawdzenie wiedzy na poziomie wyższym niż sama znajomość faktów (poziom zrozumienia, umiejętność analizy, syntezy, rozwiązywania problemów) [96].

Precyzja pomiaru, która jest miarą powtarzalności uzyskiwanych wyników w określonych warunkach (np. w sytuacji egzaminowania), stanowi o jego rzetelności. Jak pokazują przedstawione wyniki badania rzetelności, ocenianie stosowane w latach 2006/07–2013/14 charakteryzowało się niedostatecznym poziomem rzetelności, ponieważ nie osiągnięto zakładanego minimalnego progu 0,700 (współczynnik alfa-Cronbacha wyniósł 0,494). Ponadto analiza poszczególnych roczników wskazuje, że w kolejnych latach poziom wewnętrznej zgodności był zróżnicowany (alfa od 0,512 do 0,703). Całkowitą wewnętrzną zgodność pomiaru obniżały oceny egzaminacyjne studentów dla dwóch przedmiotów: podstawy psychoterapii i epidemiologia. Obserwacja ta jest więc zbieżna z ustaleniami poczynionymi w zakresie różnic w ocenianiu, które zostały opisane powyżej.

Zły dobór metod egzaminacyjnych (np. zamiast egzaminów praktycznych stosowano test), a w szczególności nieprawidłowa konstrukcja zadań sprawdzających, które są podstawą oceniania, przyczynia się do spadku rzetelności pomiaru. Student może nie mieć sposobności pokazania pełni swoich osiągnięć z danej dziedziny, jeżeli egzamin znacząco zawęży zakres treści, które służą ocenie danych efektów kształcenia. Inną przyczyną niskiej wartości współczynnika alfa może być także duży udział błędów losowych w wynikach określonego pomiaru, co ma związek z nieprawidłowym przygotowaniem lub stosowaniem punkowanego klucza odpowiedzi [68].

Trzecim istotnym parametrem stosowanym przy analizie jakości systemu oceniania jest oszacowanie trafności pomiaru. W oszacowaniu trafności kryterialnej posługujemy się analizą korelacyjną, w której badamy, w jakim stopniu części składające się na całość pomiaru tworzą silne dodatnie korelaty (*cross-correlates*) [99]. Uzyskane wyniki analizy interkorelacji są w dużym stopniu spójne z opisanymi powyżej ustaleniami dotyczącymi oceny zgodności punktowania oraz pomiarami rzetelności.

Otrzymane wyniki wskazują, że dla zdecydowanej większości przedmiotów istnieją pozytywne zależności między ocenami uzyskiwanymi przez studentów z kolejnych egzaminów. Wyjątkiem są tutaj oceny egzaminacyjne z zakresu podstaw psychoterapii, dla których średnia wartość współczynnika korelacji Spearmana osiągnęła bardzo niską wartość ($r_s = 0,02$). Brak istotnych korelacji międzyprzedmiotowych pozwala postawić tezę, że uzyskane przez studentów wyniki z podstaw psychoterapii odzwierciedlają ocenę innych treści i umiejętności zdających, niż ma to miejsce w przypadku pozostałych przedmiotów. Przypuszczenie to, chociaż mało prawdopodobne, wymaga pogłębionej analizy trafności programowej w oparciu o szczegółowe efekty kształcenia w przypadku przedmiotów, dla których brak jest pozytywnych korelatów. Bardziej prawdopodobne jest, że ocena osiągnięć studentów dla tego przedmiotu została uzyskana w wyniku źle zaplanowanego egzaminu lub że niepoprawnie przeprowadzono proces dydaktyczny. W innym przypadku należałoby podważyć trafność oceny dla pozostałych zakresów kształcenia, dla których nie wykazano pozytywnych interkorelacji z tym przedmiotem.

Ze względu na fakt, że w badaniach trafności kryterialnej przyjmowanie współczynnika korelacji Spearmana jako wskaźnika oceny bardzo upraszcza model, znacznie bardziej wiarygodne wyniki można uzyskać z wielowymiarowej analizy z użyciem metody eksploracyjnej analizy czynnikowej. Rozwiązanie otrzymane w wyniku ortogonalnej rotacji ładunków czynnikowych pozwala wnioskować, że oceny studentów uzyskiwane z poszczególnych przedmiotów nie tworzą odrębnych składowych. W szczególności dotyczy to podstaw psychoterapii, dla których stwierdzono wyraźną odrębność w zakresie ocenianych efektów kształcenia. Ustalenie to nie jest jednak zgodne z założeniami programu studiów, ponieważ podstawy psychoterapii należą do tej samej grupy przedmiotów, co zarządzanie w pielęgniarstwie, teoria pielęgniarstwa, pielęgniarstwo europejskie i dydaktyka medyczna (grupa nauk społecznych) [96]. Dodatkowo powyższa obserwacja ma także potwierdzenie w wynikach analizy głównych składowych, która to analiza wskazuje na brak jednowymiarowości skali pomiarowej oceniającej osiągnięcia studentów w zakresie nauk społecznych. Podsumowując, można stwierdzić, że w odniesieniu do trafności teoretycznej i kryterialnej otrzymane wyniki wspierają tezę o niepoprawności oceniania studentów w zakresie podstaw psychoterapii.

Niedostateczny poziom rzetelności oraz zastrzeżenia dotyczące obiektywizmu i trafności oceniania studentów na studiach drugiego stopnia skłaniają do postawienia krytycznych wniosków w zakresie wiarygodności pomiaru osiągnięć edukacyjnych. Niewystarczająca precyzja i zgodność oceniania wpływają na słabsze różnicowanie studentów, a co za tym idzie – końcowa ocena na dyplomie oraz średnia ocen z toku studiów mogą nie być dobrymi miarami jakości przygotowania

absolwenta w zakresie oczekiwanych efektów kształcenia. Ponadto niezadowalająca trafność, która jest także jednym z wyznaczników jakości systemu oceniania, sprawia, że dyplom magistra pielęgniarstwa nie gwarantuje posiadania przez absolwenta rzeczywistych kompetencji. Problem ten jest szczególnie istotny w przypadku tych studentów, dla których średnia ocen z toku studiów lokuje się w pobliżu dolnej granicy skali (średnia ocen $\approx 3,0$). Ze względu na fakt, że są to studenci najslabsi, istnieje duże ryzyko niedostatecznego poziomu kompetencji reprezentowanych przez tę grupę absolwentów. Może to skutkować spadkiem zaufania pracodawców w stosunku do osób ubiegających się o pracę i posiadających dyplom uczelni, która nie dba w należyty sposób o jakość kształcenia.

EDUKACYJNA WARTOŚĆ DODANA W OCENIE JAKOŚCI KSZTAŁCENIA

Wdrożenie mechanizmów kontroli jakości powinno uwzględniać fakt, że ocena efektywności kształcenia na danym kierunku studiów powinna być holistyczna. System oceniania powinien być spójny ze standardami kształcenia i należy dążyć do tego, aby w maksymalnym stopniu trafnie mierzył rzeczywiste kompetencje studiujących. Powyższy wymóg nakłada na uczelnię obowiązek stałego monitorowania efektywności kształcenia z zastosowaniem obiektywnych, walidowanych metod pomiaru. Szczególnie istotne jest, aby zastosowane narzędzie oceny jakości było trafne w zakresie użyteczności w pomiarze efektywności kształcenia na danej uczelni. Ze względu na brak w Polsce zewnętrznego kryterium oceny efektywności, jakie mógłby stanowić Pielęgniarski Egzamin Końcowy, prowadzony na wzór Lekarskiego lub Lekarsko-Dentystycznego Egzaminu Końcowego, konieczne jest poszukiwanie innych metod oceny trafności. Zaproponowana w niniejszym opracowaniu metoda opiera się na założeniu, że miarą efektywności kształcenia może być względny postęp edukacyjny, jaki można zaobserwować w toku kształcenia na kierunku pielęgniarstwo. Do tego celu wykorzystano metodę szacowania EWD z wyodrębnieniem dwóch poziomów:

- szczebel niższy (początkowy potencjał edukacyjny): wyniki kwalifikacji na studia oraz oceny uzyskane w toku kształcenia na studiach pierwszego stopnia,
- szczebel wyższy (zmienna wynikowa): średnia ocen dla przedmiotów egzaminacyjnych objętych programem studiów drugiego stopnia.

Wyniki analizy predykcyjnej z użyciem prostych korelacji i regresji jednozmiennowej wskazują, że generalnie sukces akademicki absolwenta studiów drugiego stopnia jest w dużej części uwarunkowany jego poziomem przygotowania do podjęcia studiów wyższych (kryteria kwalifikacji na studia pierwszego stopnia) oraz kompetencjami uzyskanymi w trakcie nauki na studiach przed uzyskaniem dyplomu licencjata. Ponadto na podstawie uzyskanych wyników dotyczących rzetelności oceniania studentów ze wszystkich przedmiotów egzaminacyjnych

na obu poziomach studiów można stwierdzić, że istnieje duży stopień zgodności w systemie oceny osiągnięć edukacyjnych. Oczywiście w tym zakresie istnieją także pewne obszary, w których należy podjąć działania projakościowe; odnosi się to w szczególności do jednego z przedmiotów z grupy nauk podstawowych.

Wyniki oceniania studentów z zakresu fizjologii wpływały niekorzystnie na rzetelność, a także były negatywnie skorelowane z efektywnością nauczania na wyższych poziomach, co świadczy o niskiej trafności tej oceny. Obserwacja ta jest o tyle niepokojąca, że gruntowna znajomość biologicznych podstaw funkcjonowania organizmu człowieka w zdrowiu i chorobie jest podstawą rozumienia przyczyn, dla których dane postępowanie kliniczne jest uważane za bezpieczne i skuteczne. Dlatego też w kształceniu przyszłych kadr pielęgniarstwa istotny nacisk kładzie się na edukację w zakresie nauk podstawowych, takich jak anatomia czy fizjologia [100]. Morag Prowse zwróciła uwagę, że kluczem do sukcesu w kształceniu kompetencji klinicznych jest dobre przygotowanie w dziedzinie nauk biologicznych [101]. W zapisach dyrektywy 2005/36/WE znajdujemy stwierdzenie, że w czasie kształcenia student pielęgniarstwa uzyskuje *odpowiednią wiedzę w zakresie nauk, na których opiera się pielęgniarstwo ogólne, łącznie z wystarczającym zrozumieniem budowy, funkcji fizjologicznych i zachowania osób zdrowych i chorych* [102]. Powyższe zalecenia wynikają z faktu, że bez pełnej wiedzy i zrozumienia anatomii czy fizjologii pielęgniarka nie jest w stanie poprawnie interpretować znaczenia poszczególnych objawów klinicznych, a tym samym podejmować adekwatnych działań [100, 101]. Wielu autorów uważa, że osiągnięte efekty kształcenia z zakresu fizjologii są, obok tych z anatomii, jednym z ważniejszych prerekwizytów późniejszego kształtowania klinicznego [100, 103, 104].

Analiza modelu predykcyjnego ujawniła, że ważnymi czynnikami wpływającymi w istotny sposób na sukces akademicki studenta na studiach drugiego stopnia były: (I) wyniki maturalne z języka polskiego oraz (II) oceny z języka obcego uzyskane w toku nauki na studiach pierwszego stopnia. Fakt ten nie jest zaskakujący, ponieważ w dostępnych danych literaturowych sporo miejsca poświęcono edukacji studentów pielęgniarstwa w obszarze kształcenia zdolności językowych i komunikacyjnych [105]. W edukacji oraz praktyce zawodowej pielęgniarki umiejętność porozumiewania się w mowie i piśmie stanowi centralny punkt [106]. Ma ona wpływ między innymi na efektywność uczenia się w trakcie kształcenia zawodowego. Susie Chen i Denise Voyles wykazały, że między uzyskaną przez kandydata na studia pielęgniarstwa pierwszego stopnia punktacją w zakresie czytania ze zrozumieniem i znajomości gramatyki a wynikami na studiach zachodzi silna korelacja [107]. Oprócz zasadniczego wpływu tej grupy umiejętności na sukcesy akademickie studenta nie mniej ważne jest także to, że stanowią one o jakości sprawowanej opieki nad pacjentem [106]. Wobec powyższego wiele

programów kształcenia przyszłych oraz aktywnych zawodowo pielęgniarek zawiera specjalne moduły przeznaczone do kształtowania kompetencji komunikacyjnych oraz językowych, w tym również takich, które są związane z umiejętnością czytania i samodzielnego pisania publikacji naukowych [105]. W związku z tym wydaje się, że uzyskane w drodze modelowania najlepiej dopasowane równanie regresji, uwzględniające wyniki kwalifikacji w zakresie języka polskiego oraz oceny egzaminacyjne z języka obcego, było jak najbardziej trafne. Oba predyktory w dużym stopniu wpływały na poziom osiągnięć absolwenta studiów drugiego stopnia, co, jak można przypuszczać, będzie miało przełożenie na jakość praktyki pielęgniarskiej.

Kolejnym ważnym predyktorem, dla którego wykazano istotny pozytywny wpływ na sukces akademicki studenta studiów drugiego stopnia, była średnia ocen z pierwszego roku studiów przeddyplomowych. Również w tym przypadku dostępne dane literaturowe potwierdzają dobre właściwości predykcyjne tego czynnika [77, 108]. Postuluje się, że wyniki kształcenia w pierwszym semestrze są wskaźnikiem tego, w jakim stopniu student przystosował się do rygoru nauki w uczelni wyższej [77].

Następnym kluczowym predyktorem, który warunkuje sukces studenta studiów drugiego stopnia i który został uwzględniony w końcowej postaci równania regresji, była punktacja z podtestu pielęgniarstwo kliniczne, będącego częścią egzaminu wstępnego. Natomiast spośród przedmiotów bloku D – grupa efektów kształcenia związanych z opieką specjalistyczną – jedynie wyniki egzaminacyjne studentów z zakresu psychiatrii spełniły założenia regresji krokowej i zostały włączone do modelu predykcyjnego. Mimo braku doniesień naukowych dotyczących podobnych badań predykcyjnych na podstawie innych przesłanek można zakładać, że opisane zależności nie są przypadkowe. Opublikowane wyniki amerykańskich badań korelacyjnych dowodzą istnienia dodatniej relacji między wynikami kształcenia na studiach przeddyplomowych (przedmioty zawodowe) a punktacją na końcowym egzaminie NCLEX-RN [109, 110]. Opierając się na cytowanych źródłach literaturowych, można wnioskować, że wyniki nauczania studentów obejmujące swoim zakresem opiekę specjalistyczną oraz przedmioty kliniczne w istotny sposób przyczyniają się do sukcesu akademickiego. Można oczywiście spekulować, na ile sukces kształcenia na studiach drugiego stopnia jest uwarunkowany podobnymi czynnikami jak powodzenie na egzaminie końcowym; jednakże wydaje się, że w sytuacji, w której 60% czasu przeznaczonego na kształcenie na studiach pierwszego stopnia jest poświęcane kształtowaniu kompetencji klinicznych, będzie to naturalny czynnik predykcyjny dla wyników nauczania na wyższym poziomie studiów. W związku z tym uwzględnienie takich zmiennych w modelu regresji przeznaczonym do szacowania EWD wydaje się bardzo uzasadnione.

Ponieważ w trakcie naboru na studia drugiego stopnia na WUM stosuje się dodatkowe kryterium punktowe za ponadstandardową aktywność naukową w trakcie studiów pierwszego stopnia, to w ramach poszukiwania najlepszego zestawu predyktorów zmienna ta również została poddana analizie. Przeglądając poszczególne składowe budujące koncepcję studiów drugiego stopnia, można zauważyć, że zdobyte przez studenta doświadczenie naukowe i umiejętności z zakresu metodologii prowadzenia badań są czynnikami sprzyjającymi zdobywaniu kluczowych kompetencji na tym poziomie studiów. Oczywiście każdy absolwent z tytułem licencjata pielęgniarstwa musiał w ramach studiów pierwszego stopnia zrealizować program z zakresu metodologii badań oraz wykonać pracę dyplomową, jednak poziom przygotowania w tym obszarze jest dalece niewystarczający.

Uzyskany w wyniku regresji krokowej model predykcyjny został przetestowany na danych zebranych z ostatniego rocznika, dla którego zakończył się pełny cykl kształcenia. Otrzymane wyniki analizy korelacyjnej wskazują, że proponowany model dobrze prognozuje przyszłe osiągnięcia studentów z okresu studiów drugiego stopnia ($r = 0,620$). Mimo że model regresji trochę gorzej radził sobie z prognozą dla studentów ze słabszymi wynikami, można uznać, że stosowany system oceniania osiągnięć studentów charakteryzował się dostateczną trafnością. Ponadto, szacując EWD dla rocznika testowego (2013/14), uzyskano wynik na poziomie 0,0067, który jednak mieści się w 95-procentowym przedziale ufności między $-0,0557$ a $+0,0691$, co oznacza, że faktycznie osiągnięcia studentów z tego rocznika nie różnią się od osiągnięć poprzedników. Ustalenia te pozwalają zakładać, że system oceniania studentów charakteryzuje się pewną stałością, ponieważ brak istotnego wzrostu lub spadku EWD świadczy o tym, że kształcenie jest utrzymywane na podobnym poziomie.

Przedstawione wyniki analizy predykcyjnej i szacowania EWD potwierdzają trafność kryterialną i prognostyczną w zakresie stosowalności i użyteczności tej metodologii w ewaluacji systemu oceniania na kierunku pielęgniarstwo na WUM. Jednakże dla pełnej oceny efektywności kształcenia konieczne jest zastosowanie metody porównawczej, która pozwoliłaby na względną ocenę osiągnięć studentów jednej uczelni wobec innej/innych. Jest to związane z brakiem referencyjnych zewnętrznych kryteriów jakościowych, z którymi można by skonfrontować efektywność kształcenia określonej instytucji akademickiej. Jak już wspomniano wcześniej, porównanie różnych uczelni jest możliwe tylko wtedy, gdy w procesie oceny osiąganych efektów kształcenia stosujemy standaryzowane narzędzia pomiaru dydaktycznego. Międzyuczelniana współpraca na tym polu mogłaby otworzyć nowe możliwości w zakresie kontroli i oceny jakości kształcenia na kierunku pielęgniarstwo, w skali większej niż tylko pojedyncza jednostka akademicka.

5.4.4. Implikacje praktyczne przeprowadzonych badań

Mimo że uzyskane w opisywanym badaniu wyniki dotyczą tylko jednej uczelni, to szereg obserwacji oraz przetestowanych rozwiązań może być użyteczny także dla innych uczelni i szkół prowadzących kształcenie na kierunku pielęgniarstwo. W oparciu o ideę Evidence-Based Admission Policy utworzono poniższe rekomendacje:

1. W celu kontroli skuteczności podejmowanych działań w obszarze polityki i strategii rekrutacyjnej uczelni należy wprowadzać kompleksowe rozwiązania w zakresie gromadzenia danych (np. socjodemograficznych, ekonomicznych, psychospołecznych) dotyczących osób kandydujących na studia oraz tych, które podjęły kształcenie na WUM.
2. Zasadniczą kwestią, która ma decydujący wpływ na ocenę jakości kształcenia na określonym kierunku studiów, jest dostęp do wiarygodnych danych. Oceny studentów uzyskiwane w toku nauki mają wartość jedynie wtedy, gdy odzwierciedlają rzeczywiste osiągnięcia edukacyjne. Dlatego tak istotne jest stosowanie i walidacja metod pomiaru dydaktycznego charakteryzujących się rzetelnością, trafnością i obiektywizmem. Wprowadzanie standaryzowanych i obiektywnych metod egzaminowania jest obecnie podstawowym zadaniem dla społeczności akademickiej.
3. Efektywność stosowanych zasad kwalifikacji na studia oraz stopień opanowania wiedzy, nabyte umiejętności i posiadane kompetencje społeczne studentów należy oceniać w perspektywie krótko- i długoterminowej. Do tego celu służyć ma dobrze zorganizowany system kompleksowej analizy wyników kształcenia (bieżąca ocena w toku studiów) oraz monitoring losów zawodowych absolwentów (wieloletnia ocena prognostyczna).
4. W celu zagwarantowania odpowiednio licznej grupy kandydatów dobrze przygotowanych do studiów pielęgniarskich pierwszego stopnia konieczne jest wdrożenie przez uczelnię rozwiązań, które będą aktywnie promowały ten kierunek studiów wśród młodzieży szkół średnich, a także w otoczeniu osób chcących podnieść swoje kwalifikacje zawodowe lub się przekwalifikować.
5. Pełna realizacja koncepcji studiów drugiego stopnia wymaga dostępu do odpowiedniej grupy kandydatów o specyficznych predyspozycjach. Z jednej strony należy więc dążyć do poprawy jakości w zakresie stosowanych zasad kwalifikacji, ale z drugiej konieczne jest także przebudowanie programu tych studiów. Realizowany program, mimo że zgodny z obowiązującymi obecnie standardami kształcenia, nie odpowiada w pełni na oczekiwania rynku pracy

w stosunku do absolwentów legitymujących się dyplomem magistra pielęgniarstwa. Atrakcyjność studiów drugiego stopnia może być kluczowa dla zapewnienia kształcenia na wysokim poziomie, które będzie docenione zarówno przez studiujących, jak i przez potencjalnych pracodawców.

5.4.5. Bibliografia

1. Słosorz T. *Zmiany akademickiego systemu kształcenia pielęgniarek i położnych w Polsce w latach 1969–2009*. Uniwersytet Medyczny im. Karola Marcinkowskiego w Poznaniu, Poznań; 2014.
2. Wstępna ocena zasobów kadrowych pielęgniarek i położnych w Polsce, do roku 2020. Naczelna Rada Pielęgniarek i Położnych, Warszawa, 2010.
<http://ww.izbapiel.org.pl/attachments/article/1782/Wstepna.ocena.zasobow.kadrowych.pdf> (data dostępu 19.06.2017).
3. Murray KT, Merriman CS, Adamson C. Use of the HESI Admission Assessment to predict student success. *Computers, Informatics, Nursing: CIN*. 2008;26:167-172.
4. Norman LP. *Prediction of nursing student performance in first year coursework*. Auburn University, 2006.
5. McNelis AM, Wellman DS, Krothe JS, Hrisomalos DD, McElveen JL, South RJ. Revision and evaluation of the Indiana University School of Nursing baccalaureate admission process. *Journal of Professional Nursing*. 2010;26:188-195.
6. Bore M, Munro D, Powis D. A comprehensive model for the selection of medical students. *Medical Teacher*. 2009;31:1066-1072.
7. Perfetto G. Predicting Academic Success in the Admissions Process: Placing an Empirical Approach in a Larger Process. *College Board Review*. 2002;196:30-35.
8. Goodwin LD. Changing conceptions of measurement validity: an update on the new standards. *Journal of Nursing Education*. 2002;41:100-106.
9. Krajowa Rada Akredytacyjna Szkół Pielęgniarek i Położnych. Uchwała nr 4/IV/2013 z dnia 21 listopada 2013 roku Krajowej Rady Akredytacyjnej Szkół Pielęgniarek i Położnych w sprawie określenia szczegółowych kryteriów oceny spełniania przez uczelnie standardów kształcenia. 2016.
10. Ben-David MF. AMEE Guide No. 18: Standard setting in student assessment. *Medical Teacher*. 2000;22:120-130.
11. Schuwirth LW, van der Vleuten CP. General overview of the theories used in assessment: AMEE Guide No. 57. *Medical Teacher*. 2011;33:783-797.
12. Sood R, Singh T. Assessment in medical education: Evolving perspectives and contemporary trends. *National Medical Journal of India*. 2012;25:357-364.
13. Niemierko B. *Diagnostyka edukacyjna*. Warszawa, Wydawnictwo Naukowe PWN; 2009.
14. Niemierko B. *Pomiar wyników kształcenia*. Warszawa, Wydawnictwo Szkolne i Pedagogiczne; 1999.
15. Norman GR, Vleuten C, Newble DI. *International handbook of research in medical education*. Dordrecht Kluwer Academic Publishers; 2002.
16. Waters A. What a waste: Nursing Standard's investigation into attrition rates from pre-registration courses produced some startling findings. *Nursing Standard*. 2006;20:14-17.

17. Banks P, Kane H, Rae C, Atkinson J. Support for nursing and midwifery students: a special case? *Nurse Education Today*. 2012;32:309-314.
18. Seago JA, Spetz J. *Admission Policies and Attrition Rates in California Community College Nursing Programs*. University of California at Berkeley California Policy Research Center; 2003.
19. Wilson R, Eva K, Lobb DK. Student attrition in the Ontario midwifery education programme. *Midwifery*. 2013;29:579-584.
20. Andrew S, Salamonson Y, Weaver R, Smith A, O'Reilly R, Taylor C. Hate the course or hate to go: semester differences in first year nursing attrition. *Nurse Education Today*. 2008;28:865-872.
21. Gillen S. Is enough being done to reduce undergraduate attrition rates? *Nursing Standard*. 2012;27:12-13.
22. Kerby MB. Toward a New Predictive Model of Student Retention in Higher Education An Application of Classical Sociological Theory. *Journal of College Student Retention: Research, Theory & Practice*. 2015;17:138-161.
23. Urwin S, Stanley R, Jones M, Gallagher A, Wainwright P, Perkins A. Understanding student nurse attrition: learning from the literature. *Nurse Education Today*. 2010;30:202-207.
24. Pitt V, Powis D, Levett-Jones T, Hunter S. Factors influencing nursing students' academic and clinical performance and attrition: an integrative literature review. *Nurse Education Today*. 2012;32:903-913.
25. Lockie NM, Burke LJ. Partnership in Learning for Utmost Success (PLUS): evaluation of a retention program for at-risk nursing students. *Journal of Nursing Education*. 1999;38:188-192.
26. Shelton EN. Faculty support and student retention. *Journal of Nursing Education*. 2003;42:68-76.
27. Shelton EN. A model of nursing student retention. *International Journal of Nursing Education Scholarship*. 2012;9: Article 6.
28. Janes S. Experiences of African-American Baccalaureate Nursing Students Examined through the Lenses of Tinto's Student Retention Theory and Astin's Student Involvement Theory. *ASHE Annual Meeting Paper*. 1997.
29. Jeffreys MR. Evaluating enrichment program study groups: academic outcomes, psychological outcomes, and variables influencing retention. *Nurse Educator*. 2001;26:142-149.
30. Jeffreys MR. Students' perceptions of variables influencing retention: a pretest and post-test approach. *Nurse Educator*. 2002;27:16-19.
31. Beil C, Reisen CA, Zea MC, Caplan RC. A longitudinal study of the effects of academic and social integration and commitment on retention. *Journal of Student Affairs Research and Practice*. 2000;37:1-10.
32. Jeffreys MR. Jeffreys's Nursing Universal Retention and Success model: Overview and action ideas for optimizing outcomes A–Z. *Nurse Education Today*. 2015;35:425-431.
33. Jeffreys MR. *Nursing student retention: Understanding the process and making a difference*. Springer Publishing Company; 2012.
34. Bean JP. Student attrition, intentions, and confidence: Interaction effects in a path model. *Research in Higher Education*. 1982;17:291-320.
35. Bean JP. Interaction effects based on class level in an explanatory model of college student dropout syndrome. *American Educational Research Journal*. 1985;22:35-64.

36. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*. 1977;84:191.
37. Grossbach A, Kuncel NR. The Predictive Validity of Nursing Admission Measures for Performance on the National Council Licensure Examination: A Meta-Analysis. *Journal of Professional Nursing*. 2011;27:124-128.
38. Dolata R, Pokropek A. Wprowadzenie. [W] Dolata R, (red.). *Edukacyjna wartość dodana jako metoda oceny efektywności nauczania na podstawie wyników egzaminów zewnętrznych*. Centralna Komisja Egzaminacyjna, Warszawa, 2007.
39. Tam M. Measuring Quality and Performance in Higher Education. *Quality in Higher Education*. 2001;7:47-54.
40. Dolata R, Pokropek A. Metody szacowania edukacyjnej wartości dodanej. [W] Jakubowski M, (red.). *Edukacyjna wartość dodana jako metoda oceny efektywności nauczania na podstawie wyników egzaminów zewnętrznych*. Centralna Komisja Egzaminacyjna, Warszawa, 2007.
41. Christensen R. *Log-linear models and logistic regression*. New York Springer Science & Business Media; 2006.
42. Wanvarie S, Sathapatayavongs B. Logistic regression analysis to predict Medical Licensing Examination of Thailand (MLET) Step1 success or failure. *Annals of the Academy of Medicine, Singapore*. 2007;36:770-773.
43. Hanley JA, Hajian-Tilaki KO. Sampling variability of nonparametric estimates of the areas under receiver operating characteristic curves: an update. *Academic Radiology*. 1997;4:49-58.
44. Cox DR. The Analysis of Multivariate Binary Data. *Journal of the Royal Statistical Society Series C (Applied Statistics)*. 1972;21:113-120.
45. Tarnowski A, Fronczyk K. Szacowanie trafności. [W] Fronczyk K, (red.). *Psychometria – podstawowe zagadnienia*. Vizja Press & IT, Warszawa, 2009; 140-160.
46. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *International Journal of Medical Education*. 2011;2:53-55.
47. Jankowski K, Zajenkowski M. Metody szacowania rzetelności pomiaru testem. [W] Fronczyk K, (red.). *Psychometria – podstawowe zagadnienia*. Vizja Press & IT, Warszawa, 2009; 84-110.
48. Bedyńska S, Cypriańska M. Zaawansowane sposoby tworzenia wskaźników: zastosowanie analizy czynnikowej oraz analizy rzetelności pozycji. [W] Bedyńska S, Brzezicka A, (red.). *Statystyczny drogowskaz. Praktyczny poradnik analizy danych w naukach społecznych*. Wydawnictwo SWPS Academica, Warszawa, 2007; 134-161.
49. Kaiser HF. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*. 1958;23:187-200.
50. Wolkowitz AA. Multiple attempts on a nursing admissions examination: Effects on the total score. *Journal of Nursing Education*. 2011;50:493-501.
51. Kulatunga-Moruzi C, Norman GR. Validity of admissions measures in predicting performance outcomes: the contribution of cognitive and non-cognitive dimensions. *Teaching and Learning in Medicine*. 2002;14:34-42.
52. Timer JE, Clauson MI. The use of selective admissions tools to predict students' success in an advanced standing baccalaureate nursing program. *Nurse Education Today*. 2011;31:601-606.
53. Janczukowicz J. Medical education in Poland. *Medical Teacher*. 2013;35:537-543.

54. Odpowiedź sekretarza stanu w Ministerstwie Edukacji Narodowej i Sportu – z upoważnienia ministra – z dnia 26 listopada 2004 r. na interpelację nr 8670 w sprawie ujednoczenia postępowania w zakresie rekrutacji na studia wyższe.
55. Dante A, Petrucci C, Lancia L. European nursing students' academic success or failure: A post-Bologna Declaration systematic review. *Nurse Education Today*. 2013;33:46-52.
56. Lancia L, Petrucci C, Giorgi F, Dante A, Cifone MG. Academic success or failure in nursing students: Results of a retrospective observational study. *Nurse Education Today*. 2013;33:1501-1505.
57. Harris MS. *Investigation of prerequisite science course performance and cumulative grade point average as predictors of success on the National Council Licensure Examination for Registered Nurses*. Tennessee State University, Tennessee, 2006.
58. Dante A, Valoppi G, Saiani L, Palese A. Factors associated with nursing students' academic success or failure: a retrospective Italian multicenter study. *Nurse Education Today*. 2011;31:59-64.
59. Salamonson Y, Andrew S, Everett B. Academic engagement and disengagement as predictors of performance in pathophysiology among nursing students. *Contemporary Nurse*. 2009;32:123-132.
60. Ali PA, Naylor PB. Association between academic and non-academic variables and academic success of diploma nursing students in Pakistan. *Nurse Education Today*. 2010;30:157-162.
61. Donaldson JH, McCallum J, Lafferty P. Can we predict successful completion of the common foundation programme at interview? *Nurse Education Today*. 2010;30:649-656.
62. Vandenhouten CL. *Predictors of success and failure on the NCLEX-RN for baccalaureate graduates*. Marquette University, Wisconsin, 2008.
63. Salamonson Y, Andrew S. Academic performance in nursing students: influence of part-time employment, age and ethnicity. *Journal of Advanced Nursing*. 2006;55:342-349; discussion 350-341.
64. van Rooyen P, Dixon A, Dixon G, Wells C. Entry criteria as predictor of performance in an undergraduate nursing degree programme. *Nurse Education Today*. 2006;26:593-600.
65. Salamonson Y, Everett B, Koch J, Andrew S, Davidson PM. English-language acculturation predicts academic performance in nursing students who speak English as a second language. *Research in Nursing and Health*. 2008;31:86-94.
66. Yin T, Burger C. Predictors of NCLEX-RN success of associate degree nursing graduates. *Nurse Educator*. 2003;28:232-236.
67. Departament Pielęgniarek i Położnych, Ministerstwo Zdrowia. *Zabezpieczenie społeczeństwa w świadczenia pielęgniarskie i położnicze – rekomendacje i ich realizacja*. Warszawa; 2015. <http://www.mz.gov.pl/wp-content/uploads/2015/09/calosc.pdf> (data dostępu 19.06.2017).
68. Niemierko B. *Testy osiągnięć szkolnych. Podstawowe pojęcia i techniki obliczeniowe*. 1 wyd. Warszawa, Wydawnictwo Szkolne i Pedagogiczne; 1975.
69. Guilford JP. *Psychometric methods*. 2nd ed. New York McGraw-Hill; 1954.
70. Kubielski W. *Podstawy pomiaru, konstruowania i ewaluacji testu dydaktycznego* Wydawnictwo Wyższej Szkoły Pedagogicznej TWP; 2006.
71. Mulholland J, Anionwu EN, Atkins R, Tappern M, Franks PJ. Diversity, attrition and transition into nursing. *Journal of Advanced Nursing*. 2008;64:49-59.
72. Pryjmachuk S, Easton K, Littlewood A. Nurse education: factors associated with attrition. *Journal of Advanced Nursing*. 2009;65:149-160.

73. Deary IJ, Watson R, Hogston R. A longitudinal cohort study of burnout and attrition in nursing students. *Journal of Advanced Nursing*. 2003;43:71-81.
74. White J, Williams WR, Green BF. Discontinuation, leaving reasons and course evaluation comments of students on the common foundation programme. *Nurse Education Today*. 1999;19:142-150.
75. Simpson R. Men in non- traditional occupations: career entry, career orientation and experience of role strain. *Gender, Work & Organization*. 2005;12:363-380.
76. Anionwu E, Mulholland J, Atkins R, Tappern M, Franks P. *Diversity, attrition and transition into nursing: The DATING project final report*. Thames Valley University, London, 2005.
77. McCarey M, Barr T, Rattray J. Predictors of academic performance in a cohort of pre-registration nursing students. *Nurse Education Today*. 2007;27:357-364.
78. Kevern J, Ricketts C, Webb C. Pre-registration diploma students: a quantitative study of entry characteristics and course outcomes. *Journal of Advanced Nursing*. 1999;30:785-795.
79. Houltram B. Entry age, entry mode and academic performance on a Project 2000 common foundation programme. *Journal of Advanced Nursing*. 1996;23:1089-1097.
80. Ofori R. Age and 'type' of domain specific entry qualifications as predictors of student nurses' performance in biological, social and behavioural sciences in nursing assessments. *Nurse Education Today*. 2000;20:298-310.
81. Jowett S, Walton I, Payne S. *Challenges and Change in Nurse Education: A Study of the Implementation of Project 2000: Executive Summary*. Slough National Foundation for Educational research in England and Wales; 1994.
82. Ehrenfeld M, Tabak N. Value of admission interviews in selecting of undergraduate nursing students. *Journal of Nursing Management*. 2000;8:101-106.
83. White J, Brownell K, Lemay JF, Lockyer JM. "What do they want me to say?" The hidden curriculum at work in the medical school selection process: a qualitative study. *BMC Medical Education*. 2012;12:17.
84. Dunn G. *Design and analysis of reliability studies: The statistical evaluation of measurement errors*. Edward Arnold Publishers; 1989.
85. Tarrant M, Ware J. Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Medical Education*. 2008;42:198-206.
86. Harasym PH, Leong EJ, Violato C, Brant R, Lorscheider FL. Cuing effect of "all of the above" on the reliability and validity of multiple-choice test items. *Evaluation & The Health Professions*. 1998;21:120-133.
87. Boland RJ, Lester NA, Williams E. Writing Multiple-Choice Questions. *Academic Psychiatry*. 2010;34:310-316.
88. Downing SM. The effects of violating standard item writing principles on tests and students: the consequences of using flawed test items on achievement examinations in medical education. *Advances in Health Sciences Education: Theory and Practice*. 2005;10:133-14.
89. Vyas R, Supe A. Multiple choice questions: A literature review on the optimal number of options. *National Medical Journal of India*. 2008;21:130-133.
90. Rodriguez MC. Three Options Are Optimal for Multiple-Choice Items: A Meta-Analysis of 80 Years of Research. *Educational Measurement: Issues and Practice*. 2005;Summer:3-13.
91. Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non-functioning distractors in multiple-choice questions: a descriptive analysis. *BMC Medical Education*. 2009;9:40.

92. Aamodt M, McShane T. A meta-analytic investigation of the effect of various test item characteristics on test scores. *Public Personnel Management*. 1992;21:151-160.
93. Haladyna TM, Downing SM, Rodriguez MC. A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*. 2002;15:309-333.
94. Cizek GJ, Robinson KL, O'Day DM. Nonfunctioning options: A closer look. *Educational and Psychological Measurement*. 1998;58:605-611.
95. Wilson T. A student selection method and predictors of success in a graduate nursing program. *Journal of Nursing Education*. 1999;38:183-187.
96. Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 9 maja 2012 r. w sprawie standardów kształcenia dla kierunków studiów: lekarskiego, lekarsko-dentystycznego, farmacji, pielęgniarstwa i położnictwa (Dz. U. 2012 nr 0 poz. 631).
97. Shirrell D. Critical thinking as a predictor of success in an associate degree nursing program. *Teaching and Learning in Nursing*. 2008;3:131-136.
98. Kendall MG. A new measure of rank correlation. *Biometrika*. 1938:81-93.
99. Meagher DG, Pan T, Wegner R, Olson AT, Overgaard SL, Mehle JJ. *PCAT Reliability and Validity*. 3rd ed. San Antonio Pearson Executive Office; 2012.
100. Smales K. Learning and applying biosciences to clinical practice in nursing. *Nursing Standard (Royal College of Nursing (Great Britain) : 1987)*. 2010;24:35-39.
101. Prowse M. Learning and using biosciences in nursing Part Two: achieving patient outcomes in perioperative practice. *Journal of Advanced Perioperative Care*. 2003;1:129-135.
102. Dyrektywa 2005/36/WE Parlamentu Europejskiego i Rady z dnia 7 września 2005 r. w sprawie uznawania kwalifikacji zawodowych.
103. Blak-Kaleta A. Międzynarodowe standardy kształcenia i kwalifikacji zawodowych pielęgniarek i położnych w Unii Europejskiej. [W] Wrońska I, Krajewska-Kulak E, (red.). *Wybrane zagadnienia z pielęgniarstwa europejskiego*. Wydawnictwo Czelej, Lublin, 2007; 63-84.
104. Phillips BC, Spurling S, Armstrong WA. *Associate Degree Nursing: Model Prerequisite Validation Study*. California Community College Associate Degree Programs by The Center for Student Success, A Health Care Initiative Sponsored Project. 2002.
105. Oermann MH, Leonardelli AK, Turner KM, Hawks SJ, Derouin AL, Hueckel RM. Systematic review of educational programs and strategies for developing students' and nurses' writing skills. *Journal of Nursing Education*. 2015;54:28-34.
106. Chant S, Jenkinson T, Randle J, Russell G. Communication skills: some problems in nursing education and practice. *Journal of Clinical Nursing*. 2002;11:12-21.
107. Chen S, Voyles D. HESI Admission Assessment Scores: Predicting Student Success. *Journal of Professional Nursing*. 2013;29:S32-S37.
108. Newton SE, Smith LH, Moore G, Magnan M. Predicting Early Academic Achievement in a Baccalaureate Nursing Program. *Journal of Professional Nursing*. 2007;23:144-149.
109. Uyehara J, Magnussen L, Itano J, Zhang S. Facilitating program and NCLEX-RN success in a generic BSN program. *Nursing Forum*. 2007;42:31-388.
110. Tipton P, Pulliam M, Beckworth C, Illich P, Griffin R, Tibbitt A. Predictors of associate degree nursing students' success students. *Southern Online Journal of Nursing Research*. 2008;8:8.

6. POZOSTAŁE OSIĄGNIĘCIA NAUKOWO-BADAWCZE

Mój łączny dorobek naukowy to **115 pełnotekstowych publikacji naukowych**, opublikowanych w czasopismach indeksowanych w bazie Journal Citation Reports® (prace posiadające punktację Impact Factor) oraz umieszczonych na liście A lub B Ministerstwa Nauki i Szkolnictwa Wyższego. W odniesieniu do **48 prac jestem pierwszym lub jedynym autorem**. W skład mojego dorobku wchodzi: 86 prac oryginalnych, 23 prace poglądowe, 5 sprawozdań oraz 1 opis przypadku.

Jestem także współautorem 2 rozdziałów w podręcznikach akademickich oraz 30 krajowych i zagranicznych rozdziałów w monografiach naukowych. Do dorobku naukowego zaliczam również autorstwo lub współautorstwo 102 streszczeń pokonferencyjnych, z których 24 są indeksowane w bazie Web of Science.

Sumaryczna wartość Impact Factor całego dorobku wynosi **21,7 punktów (884 punktów MNiSW)**. Według danych bazy Web of Science łączna **liczba cytowań** wszystkich prac to **173** (bez autocytowań), a **Indeks Hirscha** – 7.

Kolejne trzy publikacje z moim współautorstwem zostały przyjęte do druku i powinny się ukazać w pierwszej połowie 2018 roku. Ich łączny współczynnik oddziaływania IF wynosi 2,24 (45 punktów MNiSW).

Główne obszary moich zainteresowań naukowych wraz ze stosownym wykazem poszczególnych pozycji piśmiennictwa stanowiących część dorobku naukowego zamieszczono w odrębnym załączniku do wniosku o wszczęcie postępowania habilitacyjnego (Załącznik nr 5 – *Opis osiągnięć naukowo-badawczych*).

Szczegółowy, poświadczony przez Bibliotekę Główną Warszawskiego Uniwersytetu Medycznego, wykaz opublikowanych przeze mnie prac naukowych, wraz z informacją o wysokości punktacji za każdą z nich, stanowi odrębny załącznik do wniosku o wszczęcie postępowania habilitacyjnego (Załącznik nr 7 – *Analiza bibliometryczna dorobku naukowego*).

