

PRZEGLĄD SYSTEMATYCZNY/SYSTEMATIC REVIEW

Otrzymano/Submitted: 10.05.2016 • Zaakceptowano/Accepted: 02.06.2016

© Akademia Medycyny

Wykorzystanie symulacji do oceny umiejętności technicznych studentów medycyny i ratownictwa medycznego: aktualny przegląd badań***The use of simulation-based assessments of technical skills of medical and paramedic students: an up-to-date review of studies*****Mariusz Panczyk¹, Robert Gałązkowski², Joanna Gotlib¹**¹ Zakład Dydaktyki i Efektów Kształcenia, Wydział Nauki o Zdrowiu, Warszawski Uniwersytet Medyczny² Zakład Ratownictwa Medycznego, Wydział Nauki o Zdrowiu, Warszawski Uniwersytet Medyczny**Streszczenie**

Wstęp. Symulacja jest coraz powszechniej stosowana nie tylko jako narzędzie dydaktyczne, ale również jako metoda oceny praktycznych umiejętności klinicznych zarówno studentów jak i pracowników ochrony zdrowia. **Cel.** Jakie są dostępne dowody, które potwierdzałyby, że symulacja jest rzetelną i trafną metodą oceny technicznych umiejętności studentów medycyny i ratownictwa medycznego. **Materiał i metody.** Przegląd piśmiennictwa obejmował angielską literaturę naukową opublikowaną w okresie między 2000 a 2015 rokiem. Przeszukano zasoby dwóch baz bibliograficznych: Medline (*Medical Literature Analysis and Retrieval System Online*) oraz *Web of Science*. Użyto słów kluczowych, które odnosiły się do rzetelności i trafności metody symulacji stosowanej do oceniania studentów. Krytyczną analizę włączonych do przeglądu publikacji wykonano zgodnie z wytycznymi *McMasters Critical Review*, które są przeznaczone do oceny badań ilościowych. **Wyniki.** Zidentyfikowano łącznie 2199 publikacji, które odpowiadały założonym kryteriom wyszukiwania. Do szczegółowej analizy zakwalifikowano 36 pełnych tekstów, z czego ostatecznie 8 poddano krytycznej ocenie. Uzyskane wyniki wskazują, że symulacja jest lepszym narzędziem oceny technicznych umiejętności studentów w porównaniu z tradycyjnymi metodami oceniania. Dodatkowo, bardzo dobre rezultaty uzyskano, gdy symulacje użyto w kombinacji z innymi metodami oraz gdy stosowano kilka różnych scenariuszy egzaminacyjnych. Głównym ograniczeniem dostępnych dowodów była mała liczebność grup badanych, niska jakość metodologiczna oraz zawężenie badań głównie do studentów medycyny. **Wnioski.** Ocena osiągniętych efektów kształcenia z zakresu technicznych umiejętności studentów medycyny i ratownictwa medycznego może być z powodzeniem prowadzona z użyciem technik symulacyjnych. Konieczne jest kontynuowanie badań na większych i bardziej zróżnicowanych grupach studentów z uwzględnieniem ścisłych zasad metodologicznych, które pozwolą w przyszłości lepiej ocenić przydatność symulacji w ocenie wybranych umiejętności studentów. *Anestezjologia i Ratownictwo 2016; 10: 184-193.*

Słowa kluczowe: umiejętności techniczne, kompetencje, studenci medycyny, ratownicy medyczni, fizjoterapeuci

Abstract

Background. Simulation is being increasingly used not only as a teaching tool but also as a method of evaluating practical clinical skills of both healthcare students and employees. **Objective.** To present available evidence that would confirm that simulation is a reliable and valid method of assessing technical skills of medical and paramedic

students. **Material and methods.** The review of studies includes English language scientific literature published between 2000 and 2015. Resources of two bibliographic databases were searched through: Medline (Medical Literature Analysis and Retrieval System Online) and Web of Science. Key words used referred to reliability and validity of the simulation method in student assessment. The publications included in the review were critically analyzed in accordance with the guidelines of *McMasters Critical Review Form for Quantitative Studies*. **Results.** A total of 2,199 publications that met the search criteria were identified. Thirty-six full texts were qualified for a detailed analysis, of which eight were ultimately subject to critical analysis. The results indicate that simulation is a better tool to assess students' technical skills than traditional assessment methods. In addition, very good results were obtained when simulations were used in combination with other methods and when several different examination scenarios were used. The key drawback of available evidence was the small size of subject groups, low methodological quality and inclusion in the studies of mainly medical students. **Conclusions.** Simulation techniques can be successfully used to assess technical skills of medical and paramedic students. Further research is required with larger and more diversified groups of students, taking into account rigid methodological rules that will enable better assessment of the usefulness of simulation in assessing students' selected skills in the future. *Anestezjologia i Ratownictwo 2016; 10: 184-193.*

Keywords: technical skills, competences, medical students, paramedics, physiotherapists

Wprowadzenie

Ocenianie zawsze pozostawało w centrum uwagi nauczycieli i wykładowców niezależnie od etapu kształcenia i jego celu. Oceną posługujemy się we wszystkich przejawach aktywności akademickiej. Bezpośrednim celem oceniania jest monitorowanie postępów studentów w toku kształcenia, a także coraz częściej, kontrola jakości kształcenia w instytucji akademickiej [1]. Pomiar jako część oceniania dostarcza nam dokładniejszej informacji niż niekontrolowana obserwacja, ale jego zastosowanie wymaga odpowiedniego przygotowania, szczegółowych analiz *post hoc* oraz fachowej interpretacji. Dlatego też, w celu zachowania wysokiej jakości oceniania, pomiar musi być prowadzony ściśle według założonych wcześniej reguł, których poprawność funkcjonowania należy stale kontrolować [2].

W edukacji przyszłych lekarzy oraz w kształceniu innych pokrewnych zawodów, jak ratownictwo medyczne, powszechnie stosowane do oceniania są różnego rodzaju testy sprawdzające, najczęściej zbudowane na bazie serii pytań wielokrotnego wyboru z jedną opcją prawidłową (*multiple-choice questions*). Mimo że ten sposób oceniania studentów ma wiele zalet, to głównym zarzutem przeciwko stosowaniu jedynie testowej formy sprawdzianów jest brak możliwości oceny praktycznych umiejętności. W szczególności dotyczy to umiejętności technicznych, zwanych także manualnymi, dla sprawdzenia których wymagane

jest zastosowanie odmiennych metod oceniania [3]. W ostatnich dwóch dekadach na studiach medycznych znacząco wzrosło zainteresowanie nietradycyjnymi metodami oceniania umiejętności studentów. Ma to przede wszystkim związek z intensywnym rozwojem technik komputerowych oraz dostępnością do wysokiej klasy sprzętu elektronicznego. Możliwość tworzenia wirtualnego środowiska, które w dużym stopniu odwzorowuje rzeczywistość, jest głównym motorem postępu w dziedzinie symulacji medycznych [4-6]. Symulacja jest techniką stosowaną w edukacji w celu rozwijania umiejętności i doświadczenia uczących się poprzez wierne odtworzenie warunków oraz sytuacji na bazie specjalnie przygotowanych scenariuszy, opartych na rzeczywistych zdarzeniach spotykanych w praktyce klinicznej [7]. Obecnie symulacje medyczne zajmują ważne miejsce w kształceniu przeddyplomowym nie tylko przyszłych lekarzy, ale także np. pielęgniarek [8]. Symulacje zapewniają studentom bardzo dobre warunki do ćwiczenia i sprawdzania poziomu nabytych umiejętności klinicznych zarówno technicznych, jak i nietechnicznych przy równoczesnym braku ryzyka dla pacjenta. Dlatego też stosowanie symulacji medycznych ma wielu zwolenników, którzy podkreślają ważną rolę tej metody kształcenia i oceniania w okresie przed dopuszczeniem studenta do kontaktu z pacjentem.

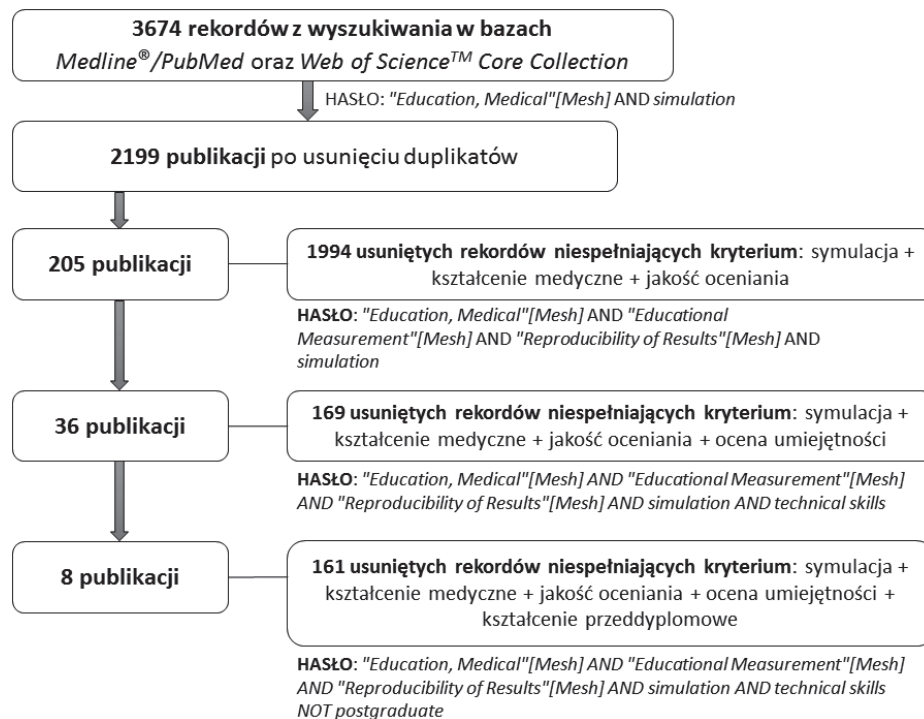
Niezbędnym elementem dobrze zaplanowanego kształcenia jest odpowiedni system oceny osiągnięć studentów [9]. Brytyjskie oraz amerykańskie wzorce

odnoszące się do jakości kształcenia medycznego (np. wytyczne *Accreditation Council for Graduate Medical Education Outcomes Project*) kładą bardzo duży nacisk na właściwy sposób oceny osiągnięć studentów w zakresie wiedzy, a przede wszystkim umiejętności zdobytych w trakcie studiów [10]. W związku z tym istnieje potrzeba opracowania dobrych narzędzi, nie tylko w zakresie nauczania, ale w równym stopniu także oceny osiąganych efektów kształcenia. Stąd równoległe z *Simulation-based Teaching* rozwijane są metody i techniki symulacyjne mające zastosowanie w pomiarze dydaktycznym (*educational measurement*), czyli tzw. *Simulation-based Assessment* (SBA) [11].

Koncepcja SBA ma swoje początki we wczesnych latach 80., kiedy to po raz pierwszy zaczęto stosować w ocenie umiejętności klinicznych egzamin typu OSCE (*Objective Structured Clinical Examination*) wykorzystujący aktorów odgrywających role pacjentów. Przystosowanie OSCE do oceny umiejętności z zakresu ratownictwa medycznego miało związek z postępowaniem technicznym, szczególnie w zakresie komputerów oraz oprogramowania. *Mannequin-based simulation* jest najczęściej stosowaną metodą w SBA, która pozwala

na odtworzenie realistycznych, a przy tym bardzo skomplikowanych, sytuacji klinicznych. Ocena umiejętności może być w tym przypadku wieloaspektowa, a informacja zwrotna udzielana studentom po egzaminie ma ogromną wartość dydaktyczną, ponieważ pozwala na zdiagnozowanie słabości w zakresie ocenianych umiejętności [12]. Te niewątpliwe zalety odróżniają SBA od tradycyjnych metod oceniania, umożliwiając obiektywną i rzetelną, a w wielu przypadkach także trafną, ocenę złożonych kompetencji, co zgodnie z przyjętymi zasadami dobrego pomiaru psychometrycznego pozwala spełnić najwyższe standardy jakościowe.

Obecnie w Polsce zachodzą duże zmiany w zakresie kształcenia lekarzy, w związku z wejściem w życie Rozporządzenia 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 poz. 631) [13]. Zgodnie z nowymi standardami kształcenia: *Sprawdzenie osiągnięcia efektów kształcenia w zakresie umiejętności praktycznych, zarówno tych, które dotyczą komuniko-*



Rycina 1. Przebieg selekcji piśmiennictwa w kolejnych etapach tworzenia listy publikacji włączonych do przeglądu
Figure 1. Selection of literature in the subsequent stages of preparing a list of publications included in the review

wania się, jak i proceduralnych (manualnych), wymaga bezpośredniej obserwacji studenta demonstrującego umiejętność w czasie tradycyjnego egzaminu klinicznego lub egzaminu standaryzowanego (OSCE) i jego modyfikacji (Mini-Cex). Egzamin OSCE jest w szczególności wskazany jako forma sprawdzania całości umiejętności klinicznych nabytych w trakcie nauczania praktycznego na ostatnim roku studiów. Ustawodawca wskazuje więc wprost na konieczność stosowania metody symulacji w ocenie umiejętności kluczowych dla takich zawodów jak lekarz czy pielęgniarka. Wzorując się na powyższych zaleceniach Ministerstwa Nauki i Szkolnictwa Wyższego można wprowadzać podobne formy oceny także w kształceniu ratowników medycznych czy fizjoterapeutów.

Mimo że stosowanie symulacji jako narzędzia dydaktycznego jest już mocno ugruntowane i w wielu aspektach dobrze opisane, to w odniesieniu do SBA nadal nie są powszechnie znane dowody, które potwierdzałyby jakość tej metody oceniania. W celu określenia, jaki jest aktualny stan wiedzy w zakresie dostępnych dowodów dotyczących skuteczności oceny umiejętności praktycznych z wykorzystaniem różnych rodzajów symulacji, zaproponowano przeprowadzenie krytycznej analizy piśmiennictwa w tym zakresie.

Materiał i metody

■ Bazy piśmiennictwa i metody wyszukiwania

Piśmiennictwo do niniejszego przeglądu wyszukiwano w dwóch bazach bibliograficzno-abstraktowych: Medline®/PubMed (*Medical Literature Analysis and Retrieval System Online*) oraz *Web of Science™ Core Collection*. Wykorzystano następujące słowa kluczowe oraz deskryptory zawarte w teaurusie pojęć medycznych MeSH (*Medical Subject Headings*): “Education, Medical”[Mesh], “Educational Measurement”[Mesh], “Reproducibility of Results”[Mesh], “simulation” oraz “technical skills”. Po usunięciu duplikatów otrzymano wynik 2199 publikacji, które odpowiadały założonym kryterium wyszukiwania. Po uwzględnieniu dodatkowych haseł zawartych w streszczeniach zredukowano początkową liczbę rekordów do 36 pozycji piśmiennictwa. Szczegółowa analiza pełnych tekstów pozwoliła na wyłonienie 8 publikacji, które zostały następnie poddane krytycznej ocenie. Schemat selekcji piśmiennictwa w kolejnych etapach tworzenia listy publikacji przedstawiono na rycinie 1.

■ Kryteria włączenia i wyłączenia

Głównym kryterium włączenia było opublikowanie pracy w języku angielskim w okresie między rokiem 2005 a 2015. Ze względu na postawiony cel niniejszego przeglądu, do analizy włączono jedynie te prace, które dotyczyły wykorzystania symulacji jako narzędzia oceny umiejętności technicznych i nietechnicznych studentów medycyny, ratownictwa medycznego i fizjoterapii. Jednakże publikacje, w których koncentrowano się na badaniu efektywności zastosowania symulacji do oceny umiejętności nietechnicznych, takich jak komunikacja, zdolności interpersonalne oraz praca zespołowa, zostały wyłączone z analizy.

Główne kryterium wyłączające dotyczyło wykorzystania symulacji jako narzędzia szkoleniowego. Ponadto wyłączono ze szczegółowej analizy wszystkie prace, których przedmiotem badania była jedynie ocena umiejętności w kształceniu podyplomowym, np. w toku rezydentury. Do przeglądu nie włączono także publikacji, które dotyczyły postaw studentów wobec sprawdzianów i egzaminów, w których wykorzystuje się metodę symulacji.

■ Krytyczna analiza

Każda z 8 włączonych do przeglądu publikacji była poddana krytycznej analizie według wytycznych *McMasters Critical Review Form for Quantitative Studies* [14]. Na skalę oceny poszczególnych badań składało się 15 pozycji punktowanych zero-jedynkowo. Wszystkie włączone do analizy badania były oceniane w oparciu o dostępne pełne teksty publikacji.

Wyniki i omówienie

Spośród analizowanej grupy publikacji, pięć dotyczyło kształcenia lekarzy, dwie ratownictwa medycznego i jedna studentów fizjoterapii. Tylko dwie prace pochodziły z ośrodków spoza Stanów Zjednoczonych (po jednej z Danii [15] i Szwajcarii [16]). Żadne badanie nie było przeprowadzone z użyciem losowego doboru próby (brak randomizacji), a także nie stosowano maskowania (brak pojedynczej lub podwójnej ślepej próby). Wszystkie badania oparte były na prospektywnej analizie obserwacyjnej, w tym trzy prace były pozbawione porównania uzyskanych wyników z grupą kontrolną. Ponadto w projektach badań nie uwzględniono statystycznych parametrów dotyczących szacowania optymalnej liczebności próby oraz nie założono *a priori* mocy statystycznej testów użytych do porównania.

Połowa spośród analizowanych badań miała charakter pilotażowy ($n < 50$), żadna praca nie miała charakteru badania wielośrodowego, co zasadniczo wpływa na brak możliwości uogólnienia uzyskanych wyników na populację ogólną. Niektóre badania uzyskały w 15-stopniowej skali oceny jakości niską punktację ze względu na brak charakterystyki grupy poddanej badaniu oraz brak jasnego opisu użytych metod do analizy statystycznej. W tabeli I zestawiono szczegółowe wyniki oceny parametrów jakościowych poszczególnych badań z użyciem skali *McMasters Critical Review Form*.

W trzech badaniach do oceny umiejętności wykorzystano symulatory średniej i wysokiej wierności

(*medium-, high-fidelity patient simulator*) [16-18]. Także w trzech innych pracach opisano wykorzystanie do symulacji metody standaryzowanego pacjenta (*standardized patient*) [19-21]. W jednej z prac Autorzy posłużyli się kombinacją dwóch metod, które do oceny wykorzystywały komputerowo wspomaganą symulację (*computer-based simulations*) w połączeniu ze standaryzowanym pacjentem [22]. Tylko jedno badanie dotyczyło zastosowania do oceniania metody symulacji opartej na środowisku wirtualnej rzeczywistości (*virtual reality*) [15]. Szczegółowe zestawienie charakterystyki poszczególnych publikacji objętych niniejszym przeglądem przedstawiono w tabeli II.

Tabela I. Wyniki oceny parametrów jakościowych dla poszczególnych badań z użyciem skali *McMasters Critical Review Form* (w przypadku pozytywnej odpowiedzi na dane pytanie przyznawano 1 punkt)

Table I. Results of the assessment of the quality parameters for the study using the *McMasters Critical Review Form* (in the case of a positive answer to the question summon 1 point)

Treść pytania	Edelstein i wsp. (2000) [22]	Boulet i wsp. (2003) [17]	Nagoshi i wsp. (2004) [19]	Grantcharov i wsp. (2005) [15]	Asprey i wsp. (2007) [20]	Panzarella i Manyon (2008) [21]	Lammers i wsp. (2009) [18]	von Wyl i wsp. (2009) [16]
Czy określono jasno cel badania?	x	x	x	x	x	x	x	x
Czy zaprezentowano wyczerpujący przegląd literatury we wstępie?	x	x	x		x	x		
Czy jednoznacznie scharakteryzowano badaną próbę?	x	x	x		x	x	x	x
Czy przedstawiono sposób szacowania liczebności próby?								
Czy rzetelnie dobrano punkty końcowe dla ocenianej interwencji?	x	x	x	x	x	x	x	x
Czy trafnie dobrano punkty końcowe dla ocenianej interwencji?	x	x	x	x	x	x	x	x
Czy szczegółowo opisano badaną interwencję?	x	x	x		x	x	x	x
Czy zastosowano metody zmniejszające ryzyko wystąpienia błędów przypadkowych i systematycznych?								
Czy opisano potencjalny wpływ dodatkowych interwencji?								
Czy uzyskane wyniki zostały poddane ocenie statystycznej?	x	x	x	x	x	x	x	x
Czy w przypadku wielu punktów końcowych, uwzględniono je w analizie statystycznej?		x				x	x	
Czy dobrano odpowiednie metody analizy statystycznej?	x	x	x	x	x	x	x	x
Czy przedstawiono możliwość zastosowania uzyskanych wyników w praktyce edukacyjnej?	x	x	x	x	x	x	x	x
Czy opisano przyczyny i liczbę nieukończonych obserwacji?								
Czy przedstawione wnioski były adekwatne do postawionego celu badania oraz użytych metod analizy?	x	x	x	x	x	x	x	x
SUMA	10	11	10	7	10	11	10	9

Tabela II. Charakterystyka poszczególnych badań włączonych do szczegółowej analizy

Table II Characteristics of studies included in the review

Autor (rok)	Charakterystyka grupy badanej	Rodzaj symulacji	Cel badania	Główne ustalenia
Edelstein i wsp. (2000) [22]	155 studentów medycyny	SP CBS	Porównanie jakości oceniania z użyciem SP i CBS względem ocen uzyskanych z egzaminów zewnętrznych	Oceny uzyskane z egzaminów z użyciem SP i CBS były zgodne z wynikami z tradycyjnego oceniania
Boulet i wsp. (2003) [17]	24 studentów medycyny 13 rezydentów	HF	Ocena rzetelności i trafności egzaminu z użyciem HF przeznaczonego dla studentów ostatniego roku studiów i pierwszego roku rezydentury	HF jest rzetelnym i trafnym narzędziem oceny umiejętności studentów i rezydentów w nagłych stanach
Nagoshi i wsp. (2004) [19]	39 studentów medycyny 49 rezydentów 11 opiekunów medycznych	SP	Ocena rzetelności i trafności egzaminu z zakresu geriatry, wyznaczenie optymalnych norm ilościowych dla tego egzaminu	Użycie SP do egzaminowania pozwala na rzetelną ocenę, jednakże Autorzy nie potwierdzili dostatecznej zdolności różnicującej tego sposobu oceny umiejętności
Grantcharov i wsp. (2005) [15]	10 studentów medycyny 10 rezydentów 8 lekarzy	VR	Ocena trafności egzaminu wykorzystującego VR do oceny umiejętności wykonania endoskopowego badania przewodu pokarmowego (badanie porównawcze trzech grup)	Potwierdzono trafność treściową (konstrukcyjną) egzaminu z użyciem VR
Asprey i wsp. (2007) [20]	101 studentów medycyny 71 studentów asystentów lekarza	SP	Analiza porównawcza wyników egzaminu dla studentów trzeciego roku z dwóch kierunków studiów (ocena umiejętności klinicznych)	Egzamin oparty na wykorzystaniu SP nie różnicuje poziomu umiejętności między dwoma grupami badanych studentów
Panzarella i Manyon (2008) [21]	34 studentów fizjoterapii	SP	Ewaluacja wielostopniowego egzaminu końcowego przeznaczonego do oceny umiejętności z zakresu fizykoterapii	Potwierdzono trafność treściową egzaminu z użyciem SP oraz rzetelność oceny, ale doskonalenie egzaminu wymaga dalszej ewaluacji
Lammers i wsp. (2009) [18]	212 absolwentów ratownictwa medycznego	LF, MF, HF	Identyfikacja braków w zakresie umiejętności praktycznych ratowników medycznych związanych z udzielaniem pomocy w nagłych wypadkach w pediatrii	Użycie symulatorów o różnym stopniu wierności pozwoliło na identyfikację braków w zakresie wybranych umiejętności praktycznych
von Wyl i wsp. (2009) [16]	30 absolwentów ratownictwa medycznego	MF	Ocena umiejętności technicznych i nietechnicznych ratowników medycznych podczas okresowych szkoleń doskonalących	Zarówno umiejętności techniczne jak i nietechniczne mogą być oceniane z użyciem symulatora pacjenta, stwierdzono wysoki stopień korelacji między oceną z obu zakresów umiejętności

SP – standaryzowany pacjent (*standardized patient*), CBS – symulacja wspierana komputerowo (*computer-based simulations*), HF – symulator pacjenta wysokiej wierności (*high-fidelity patient simulator*), MF – symulator pacjenta średniej wierności (*medium-fidelity patient simulator*), LF – symulator pacjenta niskiej wierności (*low-fidelity patient simulator*), VR – wirtualna rzeczywistość (*virtual reality*)

▪ **Symulatory o różnym stopniu wierności**

Pośród trzech badań wykorzystujących do oceny symulatory, jedno dotyczyło przypadku użycia symulatora pacjenta zaprojektowanego na potrzeby wykonywania inwazyjnych zabiegów, które nie wymagały wysokiego stopnia wierności. W badaniu von Wyl i wsp. (2009) [16] oceniano zarówno umiejętności techniczne, jak i nietechniczne w dwóch scenariuszach (ostry zespół wieńcowy oraz ciężki uraz mózgu), dla których były opracowane specjalne standaryzowane listy kontrolne z odpowiednią punktacją (tzw. *checklists*). Poprawność wykonanych czynności była oceniana na podstawie rejestracji wideo przez dwóch niezależnych ekspertów: lekarza specjalisty w zakresie medycyny ratunkowej oraz psychologa. Korelacja między ocenami wystawionymi przez obu egzaminatorów była wysoka, co skłoniło Autorów do postawienia wniosków o dobrej rzetelności takiego sposobu oceniania. Ponadto badacze zanotowali pozytywną relację między poziomem umiejętności technicznych i nietechnicznych [16].

Drugim analizowanym przypadkiem dotyczącym wykorzystania symulatora było badanie Boulet i wsp. (2003) [17], w którym użyto symulatora wysokiej wierności (*life-size simulator* MEDSIM-EAGLE® Med-Sim USA, Inc.). Autorzy przeanalizowali wyniki egzaminowania studentów ostatniego roku studiów lekarskich i pierwszego roku rezydentury w zakresie 10 różnych scenariuszy klinicznych. Szacowanie rzetelności oceniania za pomocą wyznaczenia indeksów dyskryminacyjnych wykazało, że wszystkie użyte scenariusze charakteryzowały się wysoką zdolnością różnicowania, zarówno w grupie studentów, jak i rezydentów. Ponadto Autorzy wykazali, że wylosowanie określonego scenariusza przez zdającego na egzaminie nie miało istotnego wpływu na powtarzalność oceniania, co potwierdzało wysoką rzetelność oceniania umiejętności. W zakresie oceny trafności prognostycznej egzaminu, Boulet i wsp. (2003) [17] stwierdzili, że istnieje istotna zależność między czasem spędzonym przez studenta na zajęciach praktycznych (kardiologia, intensywne terapia, medycyna ratunkowa oraz chirurgia) a sumarycznym wynikiem punktowym na egzaminie z użyciem symulatora wysokiej wierności. Dodatkowo z analizy porównawczej wyników egzaminu studentów i rezydentów wynikało, że pierwsza grupa osiągnęła statystycznie niższe oceny niż druga grupa (57,1 vs. 64,9; $p < 0,01$). Wynik ten potwierdził trafność diagnostyczną egzaminu [17].

Ostatnim z analizowanych badań związanych z użyciem symulatora była praca Lammers i wsp. (2009) [18]. Autorzy opisali ocenę wybranych umiejętności technicznych absolwentów ratownictwa medycznego w celu identyfikacji braków w zakresie udzielenia pomocy dzieciom. W procesie oceniania wykorzystano kilka typów symulatorów o różnym stopniu wierności. Zastosowane scenariusze odnosiły się do sytuacji klinicznych spotykanych w praktyce pediatrycznej: zatrzymanie krążenia i oddychania u niemowląt (*low-fidelity training manikin* LAERDAL ALS Baby, Laerdal Corp.), duszność związana z astmą u dzieci (*medium-fidelity patient simulator*, MEGACODE Kid, Laerdal Corp.) oraz wstrząs septyczny u 6-miesięcznego niemowlęcia (*high-fidelity infant simulator*, SIMBABY, Laerdal Corp.) [18]. W celu przeprowadzenia analizy jakości oceniania, zdający zostali podzieleni na dwie podgrupy różniące się poziomem doświadczenia zawodowego oraz stażem pracy. Dla wszystkich trzech scenariuszy Autorzy wykazali istotne różnice w uzyskanej punktacji między dwiema porównywanymi podgrupami [18]. Natomiast nie wykazano istotnej korelacji między liczbą lat stażu pracy a uzyskanym wynikiem z egzaminu, niezależnie od użytego scenariusza. W podsumowaniu badacze podkreślili, że symulatory o różnym stopniu wierności nie muszą różnić się w zakresie jakości oceniania umiejętności technicznych. Wybór rodzaju użytego symulatora powinien być uzależniony od scenariusza i celu prowadzenia egzaminu. W wielu sytuacjach można z powodzeniem zastosować symulator niskiej wierności, nie tracąc przy tym na jakości oceniania [18].

▪ **Standaryzowany pacjent**

Połowa spośród 8 poddanych analizie badań w niniejszym przeglądzie wykorzystywała do oceniania symulację opartą na standaryzowanym pacjencie (SP) [19-22]. Odpowiednio wyszkoleni aktorzy są najczęściej angażowani do egzaminów typu OSCE, które w wielu krajach stanowią standard w uczelniach medycznych. Poszczególne badania różniły się łącznym czasem trwania egzaminu, liczbą stacji i czasem przeznaczonym na ich zaliczenie, liczbą SP oraz rodzajem scenariuszy i zakresem ocenianych umiejętności. Zastosowanie w trakcie egzaminu symulacji opartej wyłącznie na SP spotykane jest znacznie rzadziej niż typowy OSCE. Analizowane badania charakteryzowały się znacznie dłuższym czasem, jaki musiał spędzić zdający na stacji z SP, niż ma to miejsce

przy stacjach w egzaminach OSCE. Studenci musieli wykazać się znajomością od kilku do nawet kilkunastu umiejętności, które składały się na całościowy wywiad oraz zastosowanie wybranej terapii [21]. Taki sposób egzaminowania jest zatem odmienny od tego, który spotykamy w OSCE, gdzie na danej stacji sprawdzane są z reguły pojedyncze umiejętności, a każda stacja rzadko obejmuje ocenę złożonych czynności.

Do sprawdzenia umiejętności zdających posłużono się, w zależności od badania, oceną wystawianą przez nauczyciela akademickiego [19-22] i doświadczonego klinicystę [21], a także samooceną studenta [21] i oceną przyznaną przez osobę odgrywającą rolę pacjenta [19,21]. We wszystkich analizowanych badaniach egzaminatorzy byli poddawani odpowiedniemu szkoleniu w zakresie oceniania, ale tylko w jednym badaniu oszacowano rzetelność wystawianej punktacji w trakcie przebiegu symulacji [21]. Panzarella i Manyon (2008) [21] wykazali, że istnieje silna zależność między punktacją przyznaną przez nauczyciela akademickiego a tą, którą wystawił doświadczony klinicysta. Natomiast w tym samym badaniu nie wykazano istotnej korelacji z oceną przyznaną przez aktora odgrywającego rolę pacjenta podczas egzaminu [21]. Dodatkowo w badaniu opublikowanym przez Edelstein i wsp. (2000) [22] zastosowano kombinację dwóch ocen. Pierwsza była wystawiana na podstawie przebiegu symulacji z użyciem SP, a druga dotyczyła symulacji wspieranej komputerowo (*computer-based simulations*). W przypadku studentów, którzy zaliczyli oba rodzaje egzaminów, stopień korelacji uzyskanych wyników był na poziomie $r = 0,40$, podczas gdy w grupie, która nie zaliczyła co najmniej jednego z egzaminów, zgodność wyników była wyraźnie niższa ($r = 0,24$) [22].

■ Wirtualna rzeczywistość

Mimo że techniki wirtualnej rzeczywistości (VR) są coraz częściej stosowanym narzędziem w symulacjach medycznych, to odnaleziono tylko jedno badanie, które dotyczyło oceniania z użyciem tego narzędzia. Grantcharov i wsp. (2005) [15], w celu oszacowania trafności egzaminu wykorzystującego symulację VR (GI MENTOR II *computer system*, Symbionix Ltd.), przeprowadzili analizę porównawczą trzech grup zdających: studentów medycyny, rezydentów oraz doświadczanych lekarzy. Scenariusz obejmował ocenę umiejętności w zakresie poprawnego wykonania badania endoskopowego przewodu pokarmowego. Na

podstawie wyników uzyskanych przez uczestników egzaminu z trzech kolejnych grup, Autorzy ustalili, że użycie symulacji VR w dostatecznym stopniu różnicowało poziom umiejętności zdających [15]. Na tej podstawie można wnioskować, że użyty symulator VR jest skutecznym narzędziem w odróżnianiu doświadczonych i początkujących endoskopistów. Konieczne jest jednak dalsze badanie w zakresie trafności prognostycznej, które oceni na ile dobry wynik uzyskany przez zdającego w warunkach symulowanych przekłada się na poprawne wykonanie endoskopii w warunkach klinicznych z udziałem pacjenta.

Podsumowanie

Zaprezentowana analiza dostępnych publikacji dotyczących wykorzystania symulacji medycznych w ocenie umiejętności technicznych studentów medycyny, ratownictwa medycznego i fizjoterapii wskazuje, że w szczególności symulatory różnego stopnia wierności oraz stosowanie SP dają rzetelną i trafną ocenę zdających. Uzyskane w poszczególnych badaniach wyniki generalnie potwierdzają zasadność stosowania narzędzi symulacyjnych w kształceniu przyszłych lekarzy oraz przedstawicieli innych zawodów medycznych. Należy jednak zaznaczyć, że kluczowym czynnikiem decydującym o jakości oceniania jest właściwe opracowanie scenariusza i adekwatne dobranie metody symulacji.

Obecnie dostępne dowody nie dają podstaw do wnioskowania na temat bardzo ważnego aspektu oceniania, a mianowicie szacowania stopnia trafności prognostycznej zastosowanych metod egzaminowania. Ostatecznie bowiem najważniejsze jest czy uzyskany wynik z egzaminu przeprowadzonego w trakcie kształcenia przeddyplomowego istotnie koreluje z późniejszym poziomem umiejętności, które absolwent reprezentuje podczas praktyki „przy łóżku pacjenta”. Ten obszar badań, jak na razie, nie został w dostatecznym stopniu poddany ocenie przez badaczy zajmujących się symulacjami medycznymi. Dowody wspierające tezę o zgodności ocen uzyskanych podczas symulacji z faktycznie posiadanymi umiejętnościami mogą być bardzo istotne w sytuacji, kiedy wynik takiego egzaminu decyduje o dopuszczeniu zdającego do praktyki klinicznej. Bez pewności w tym zakresie istnieje potencjalne ryzyko, że stażyści oraz rezydenci nie będą w stanie zapewnić swoim pacjentom bezpiecznej i skutecznej terapii czy opieki.

Oprócz ograniczeń metodologicznych opisach wstępnie do omawiania wyników, należy także odnieść się do problemu braku w ocenianych pracach dobrych protokołów dla analizy porównawczej. Tylko jedno badanie dotyczyło oceniania umiejętności studentów za pomocą symulacji w porównaniu z inną formą oceny lub innym rodzajem symulacji [22]. Zaplanowanie badania opartego o analizę porównawczą pozwala lepiej określić różnice między odmiennymi formami oceniania, także w obrębie różnych typów symulacji. Wnioski z tego rodzaju badań są szczególnie cenne dla nauczycieli akademickich, ponieważ umożliwiają wybór najlepszych sposobów egzaminowania w określonych warunkach technicznych i finansowych uczelni.

Dalsze badania w zakresie użycia symulacji do oceniania studentów wymagają większego niż dotychczas rygoru metodologicznego. Szczególnie dotyczy to kontroli wpływu błędów przypadkowych i systematycznych na uzyskiwane wyniki oraz wprowadzenie do protokołu badania ślepej próby. Ponadto, ze względu na praktyczny charakter tych badań, bardzo istotne jest zamieszczanie szczegółowego opisu wprowadzanej interwencji, tak aby inni zainteresowani tematem mogli w sposób jednoznaczny odtworzyć warunki badania w swoich ośrodkach akademickich. Ponieważ badania skuteczności kształcenia są narażone na liczne wypaczenia związane z wpływem różnych czynników zakłócających, to szczególnie ważne jest opracowanie dobrej i przemyślanej ich konstrukcji.

Ograniczenia

Istotnym ograniczeniem prezentowanego przeglądu jest objęcie analizą wyłącznie literatury anglojęzycznej. Nie jest wykluczone, że w zakresie omawianym w niniejszej pracy, dostępne są także dodatkowe wyniki badań opublikowane w innych językach, które mogłyby pogłębić naszą wiedzę dotyczącą skuteczności stosowania SBA. Ponadto, ze względu na bardzo zróżnicowany

charakter poszczególnych badań objętych przeglądem, nie było możliwości przedstawienia ilościowej syntezy danych, co mogłoby znacznie wzmocnić siłę dowodów potwierdzających tezę o dobrej jakości egzaminowania z użyciem symulacji. W końcu nie przedstawiono także danych, na podstawie których można by dokonać spójnej oceny efektywności kosztowej SBA, co jest bardzo istotnym kryterium branym pod uwagę przez uczelnie wdrażające w praktyce tego rodzaju rozwiązania.

Słownik skrótów:

CBS – symulacja wspierana komputerowo (*computer-based simulations*)

HF – symulator pacjenta wysokiej wierności (*high-fidelity patient simulator*)

LF – symulator pacjenta niskiej wierności (*low-fidelity patient simulator*)

MeSH – tezaurus terminów medycznych (*Medical Subject Headings*)

MF – symulator pacjenta średniej wierności (*medium-fidelity patient simulator*)

OSCE – Obiektywny Strukturyzowany Egzamin Kliniczny (*Objective Structured Clinical Examination*)

SBA – ocenianie z wykorzystaniem symulacji (*Simulation-based Assessment*)

SP – standaryzowany pacjent (*standardized patient*)

VR – wirtualna rzeczywistość (*virtual reality*)

Konflikt interesów / Conflict of interest

Brak/None

Adres do korespondencji:

✉ Mariusz Panczyk

Zakład Dydaktyki i Efektów Kształcenia

Wydział Nauki o Zdrowiu, WUM

ul. Żwirki i Wigury 81; 02-091 Warszawa

☎ (+48 22) 572 04 90

📧 mariusz.panczyk@wum.edu.pl

Piśmienictwo

1. Astin AW. Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education. Lanham: Rowman & Littlefield Publishers; 2012.
2. Niemierko B. Diagnostyka edukacyjna. Warszawa: Wydawnictwo Naukowe PWN; 2009.
3. Epstein RM. Medical education - Assessment in medical education. N Engl J Med. 2007;356:387-96.
4. Issenberg SB, Scalese RJ. Simulation in health care education. Perspect Biol Med. 2008;51:31-46.

5. Khan K, Pattison T, Sherwood M. Simulation in medical education. *Med Teach*. 2011;33:1-3.
6. Weller JM, Nestel D, Marshall SD, Brooks PM, Conn JJ. Simulation in clinical teaching and learning. *Med J Aust*. 2012;196:594.
7. Gaba DM. The future vision of simulation in healthcare. *Simul Healthc*. 2007;2:126-35.
8. Alinier G, Hunt WB, Gordon R. Determining the value of simulation in nurse education: study design and initial results. *Nurse Educ Pract*. 2004;4:200-7.
9. Duszyński M. Efekty kształcenia w Polsce: perspektywa brytyjska. *Nauka*. 2011;1:137-44.
10. Swing SR. The ACGME outcome project: retrospective and prospective. *Med Teach*. 2007;29:648-54.
11. McLaughlin S, Fitch MT, Goyal DG, Hayden E, Kauh CY, Laack TA, et al. Simulation in graduate medical education 2008: a review for emergency medicine. *Acad Emerg Med*. 2008;15:1117-29.
12. Toolbox of Assessment Methods [<http://njms.rutgers.edu/culweb/medical/documents/ToolboxofAssessmentMethods.pdf>] (dostęp 31.05.2016)
13. 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). In.
14. Law M, Stewart D, Pollock N, Letts L, Bosch J, Westmorland M. Critical review form - Quantitative studies. McMaster University: Occupational Therapy Evidence-Based Practice Research Group 1998.
15. Grantcharov TP, Carstensen L, Schulze S. Objective assessment of gastrointestinal endoscopy skills using a virtual reality simulator. *JSLs*. 2005;9:130-3.
16. von Wyl T, Zuercher M, Amsler F, Walter B, Ummenhofer W. Technical and non-technical skills can be reliably assessed during paramedic simulation training. *Acta Anaesthesiol Scand*. 2009;53:121-7.
17. Boulet JR, Murray D, Kras J, Woodhouse J, McAllister J, Ziv A. Reliability and validity of a simulation-based acute care skills assessment for medical students and residents. *Anesthesiology*. 2003;99:1270-80.
18. Lammers RL, Byrwa MJ, Fales WD, Hale RA. Simulation-based assessment of paramedic pediatric resuscitation skills. *Prehosp Emerg Care*. 2009;13:345-56.
19. Nagoshi M, Williams S, Kasuya R, Sakai D, Masaki K, Blanchette PL. Using standardized patients to assess the geriatrics medicine skills of medical students, internal medicine residents, and geriatrics medicine fellows. *Acad Med*. 2004;79:698-702.
20. Asprey DP, Hegmann TE, Bergus GR. Comparison of Medical Student and Physician Assistant Student Performance on Standardized-Patient Assessments. *J Physician Assist Educ*. 2007;18:16-9.
21. Panzarella KJ, Manyon AT. Using the integrated standardized patient examination to assess clinical competence in physical therapist students. *J Phys Ther Educ*. 2008;22:24.
22. Edelstein RA, Reid HM, Usatine R, Wilkes MS. A comparative study of measures to evaluate medical students' performance. *Acad Med*. 2000;75:825-33.